Australian Government
Department of Defence
Defence Science and
Technology Organisation

Review of Methods and Approaches for the Structural Risk Assessment of Aircraft

Paul White

Air Vehicles Division

Defence Science and Technology Organisation

DSTO–TR–1916

ABSTRACT

This report looks at the published literature on methods and assumptions made in performing structural risk assessments on aircraft. Because the major contributor to the risk of structural failure is fatigue, most methods of risk assessment involve modelling the effect of fatigue growth by some probabilistic method. Many risk assessments use the equivalent initial flaw size approach to allow for the variability in fatigue crack growth. Common errors in the formulation are made in many risk assessments, which can be significant and are described in this report. It is found that the standard approach can produce an acceptable assessment of the probability of failure of an aircraft if care is taken in understanding what is being modelled and the assumptions on which the analysis is based. A number of case studies of risk assessment's performed on different aircraft are summarised.

Published by

Defence Science and Technology Organisation
506 Lorimer St,
Fishermans Bend, Victoria 3207, Australia

Telephone: (03) 9626 7000
Facsimile: (03) 9626 7999

AR No. 013–746
October, 2006

APPROVED FOR PUBLIC RELEASE

Review of Methods and Approaches for the Structural Risk Assessment of Aircraft

EXECUTIVE SUMMARY

Analysis of the risk of failure of airframes has seen increased international activity over the past decade. It provides an alternative to the airworthiness standards for maintaining safe operation of ageing aircraft. DSTO has been developing a risk capability in recent years, and seeks to learn as much as possible from international developments and applications.

This report conducts a review of the methods and assumptions made in structural risk assessments. Its aim is to provide a guide for conducting a structural risk assessment of an airframe, and as such is limited to techniques generally used in the aircraft industry.

The literature was reviewed to investigate how the risk of fatigue failure for an aircraft has previously been determined. Factors include the initial defects that trigger fatigue cracks, the growth of fatigue cracks to failure, the determination of residual strength and the variability in the spectrum loads and the maximum loads that result in failure.

Towards the end of the life of an aircraft, the overwhelming risk or probability of failure is caused by the growth of fatigue cracks in the aircraft structure from the loads sustained through usage of the aircraft throughout its lifetime. A summary of the fatigue crack growth laws used in probabilistic risk assessments is given.

Guidance is given on what distributions should be used for modeling the various parameters in a probabilistic risk assessment. Data and a knowledge of the mechanisms that produce the variability in each parameter will define the probabilistic distributions that should be used.

The most common error found in many of the risk assessments examined is to assume the risk of failure is simply the summation of the individual risks of failure of all the elements that can fail in the aircraft. This is incorrect. The most critical crack will fail first and the risk of failure is simply the probability of the largest crack failing. The remaining cracks in the aircraft do not directly contribute to the risk of failure. This difference is significant when there is a large number of possible fasteners that can initiate failure, such as occur in the wing of transport aircraft. The number of cracks present contribute indirectly by changing the probability distribution of the size of the largest crack.

Summaries of a number of risk assessments on different aircraft are provided, such as for the KC135, F-16, C-141, B1-B Bomber, F/A-18 and the B707. These assessments have been used to support keeping the associated aircraft in service. DSTO has learned from these risk assessments in order to apply the capability to ADF aircraft, thereby evaluating and maintaining their safety of flight as they age.

Contents

1 Introduction

This report conducts a general review of the methods and assumptions made in structural risk assessments. Its aim is to establish a guide for conducting a structural risk assessment of an airframe, and as such is limited to techniques generally used in the aircraft industry.

The literature was reviewed to investigate how the risk of fatigue failure for an aircraft has previously been determined. Factors include the initial defects that trigger fatigue cracks, the growth of fatigue cracks to failure, the determination of residual strength and the variability in the spectrum loads and the maximum loads that result in failure.

Towards the end of the life of an aircraft, the overwhelming risk or probability of failure is caused by the growth of fatigue cracks in the aircraft structure from the loads sustained through usage of the aircraft throughout its lifetime[1]. These fatigue cracks may begin either from the initial defects in the material such as porosity or inclusions, or from methods of manufacture and construction. Cracks can also start from corrosion developed later in the life of the aircraft.

The risk of failure may be reduced by carrying out inspections on the aircraft (to detect and remove fatigue cracks), or from setting a conservative life limit for the aircraft, below which there is little chance of failure from fatigue. When and how often the inspections are carried out, and what technique is used, will also affect the number of cracks and size that are found and hence influence the risk of failure.

In addition, the method used to repair the aircraft may also impact on the risk. While simple repairs such as removing small cracks by oversizing a hole and fastener are generally safe, more complex repairs may not have the life expected [2] and have resulted in the loss of aircraft such as the Japan Airlines 747 [3]. Thus the risk assessment of repairs is even more difficult than a risk assessment of the aircraft itself because of the general lack of experimental data, mostly due to the one-off nature of repairs.

Methods used to set safe lives or inspection intervals are generally based on one or more of the internationally recognised airworthiness standards such as DEFSTAN 970 [4] or the Joint Services Specification Guide (JSSG) [5] which superseded the standards MIL-A-8860B [6] used by the U.S. Navy (USN) and MIL-A-87221 [7] used by the U.S. Air Force (USAF). These standards generally have some reference to safety based on acceptable probabilities of failure obtained either by a probabilistic approach or empirically through the use of scatter factors [8]. These acceptable limits are based on the cumulative (also known as total) probability of failure (CPOF or TPOF), or the instantaneous or Single Flight Probability of Failure (SFPOF). DEFSTAN 970 specifies safety limits in terms of a total probability of failure of 1/1000 for an aircraft, whereas the JSSG specifies acceptable levels of $1/10^7$ per flight hour.

A major contribution to the analysis of the risk of failure was the development of the Equivalent Initial Flaw Size (EIFS) approach by Rudd and Gray [9][10][11] in the late 1970s. This was an effective method to describe the distribution of flaws at the start of

[1]Risk can be defined as the probability of failure times the consequence of failure [1]. However in relation to aircraft where the consequence of failure is generally high, the risk is usually defined just as the probability of failure.

fatigue growth. Since the prediction of fatigue crack growth rates is in general somewhat different from the measured crack growth rates, the EIFS is required to be used with the predicted rate of crack growth to produce the correct crack size at some later date. Thus the EIFS will in general be different from a directly measured initial flaw size and is model dependent.

If the initial flaw shape is irregular, this also makes it difficult for a single size measurement to characterise the flaw. However, after a small period of growth the shape of the crack should develop and generally becomes more regular. Thus, by back extrapolating fractographically measured crack growth from the regular shaped crack, one can obtain an effective regular initial flaw size. Researchers at DSTO have called this the Equivalent Pre-crack Size (EPS) to differentiate it from EIFS which is usually back extrapolated using analytical crack growth curves instead of from fractographic data (although in some instances it has also been obtained using fractography).

EIFS is defined by Manning and Yang [12] as an artificial crack size which results in an actual crack size at an actual point in time when the initial flaw is grown forward. Implicit in this approach is that it must be grown forward by the same means as was used to obtain the EIFS. However when different load spectra are used, caution must be exercised, since it is assumed that the EIFS is entirely independent of the load spectrum. This is not always the case, particularly in instances where the high loads result in cracking of inclusions which can then lead to a fatigue crack, which would otherwise not occur, or where high loads induce other non-linear behavior such as extensive plasticity.

Another approach used is to grow cracks back from a time to failure distribution [13]. However, this has the significant drawback that it cannot be independent of the loading sequence and thus has to be re-derived for each new application. EPS is perhaps closer to a material/manufacturing process property.

The growth of fatigue cracks in aluminium components usually occur when cyclic loads are applied to the structure. Cracks usually initiate from material or manufacturing defects or processes. Other materials may exhibit more delay in the initiation phase of crack growth.

1.1 The basic steps followed in a risk analysis

A risk analysis works out the probability of failure due to variation in factors such as crack size and loads. These variables can be characterised by random distributions which characterise all things unknown about the particular variable assuming independence from other variables.

1. Risk can be calculated using probabilistic fracture mechanics. Which means it is necessary to understand the initial flaws, the crack growth, the residual strength with cracks of different sizes, and the loads which can affect the crack growth and the residual strength.

2. Generally the crack that fails is the largest[2] Thus you only need to analyze the distribution of the largest crack not of all cracks. However, the more critical locations

[2]Actually the concept of the largest cracks assumes all locations are at the same stress level. This can

that can potentially fail, the more likely an aircraft will contain an extreme sized crack, and thus the larger the critical crack is likely to be. The distribution of the largest crack can be obtained from knowing the distribution of individual cracks using mathematical extreme value theory.

3. The probability of failure is obtained by combining all the probabilities of having a configuration that can fail with the probability of experiencing a load greater than the failure load in the time interval under consideration (providing they are all independent).

4. The probability of failure can be calculated using the following methods:

   (a) simple Monte Carlo approach (random number simulation) or more complex Monte Carlo such as importance sampling which targets specified areas of the distributions (typically the tails) for sampling. This restricted targeting is then compensated for in the analysis,

   (b) numerical integration to integrate the probability of each configuration failing, and

   (c) transformations of the strength/load distributions into approximate distributions that can more easily be solved by methods such as First (or Second) Order Reliability Methods (FORM/SORM). This is particularly effective when there a many random variables to be included in the analysis. Generally this involves transformations of the tails of the distributions of interest, into normal distributions where the overlap in the tails determines the probability of failure. The distribution is so chosen that the values of the equivalent distributions produce an accurate representation in the region of the overlap, but may be significantly different elsewhere. In practice for aircraft it seems that only a few sources of scatter dominate the final answer, so the remaining sources can be fixed. This then means the previous two approaches are better suited for structural risk assessments.

5. Many distributions are available to model the random variables. Some distributions are based on some underlying processes that may be present. They are:

   (a) normal - cumulative process of random amounts

   (b) lognormal - multiplicative process of random amounts

   (c) exponential - random failures

   (d) gamma - multiple random failures

   (e) Gumbel, Frechet and reversed Weibull - extreme value (first failure in a group of identical items)

6. You can tell if the data follows a particular distribution by plotting it on 'probability paper'. These days, using computer methods (instead of manually plotting the data on paper) involves transforming the data, so that if it is from the targeted distribution, a plot of probability exceedances will produce a straight line.

be generalised to the most critical crack if we consider the extra variable of stress. The cracks can be expressed in terms of a new parameter such as criticality.

7. The variables in order of importance to the risk of failure from fatigue in aluminium aircraft tend to be [14]:

   (a) initial flaw size

   (b) loads

   (c) crack growth variability

   (d) material variability

   For a transport aircraft, for which the high loads are experienced rarely, the loads variable is more important than for a fighter aircraft, for which the high loads are experienced in virtually every flight.

8. Probability of failure can be the total probability of failure or the instantaneous probability of failure or hazard. Typical criteria for each are 10^{-3} for total probability of failure and 10^{-7} for single flight probability of failure.

9. Because the cracks grow with time, so the probability of failure usually increases with time (except for the beneficial effect of inspections).

2 History

Reliability theory was first applied by Weibull to structural materials where he developed a new distribution which modeled the strength of material based on the weakest link to fail approach. This was applied to static strength of materials, but in 1949 [15] he showed it was also applicable to time to failure due to fatigue.

The use of probabilistic risk assessments in structural reliability was first used on aircraft by Alfred M. Freudenthal in 1953 [16] and later in 1967 [17]. Whilst working at Wright Patterson Air Force Base in the United States during the 1960s, Freudenthal was the first to apply probabilistic methods to the problem of scatter in aircraft fatigue lives. He looked at the probability that a random load would produce a stress that exceeded a randomly distributed strength. He expressed his answer in terms of the time to first failure in the fleet. The larger the fleet the greater the expected probability of failure.

Working for him on attachment from ARL (now DSTO) was Alf Payne. Together they published a number of papers on the early application of probabilistic methods to the fatigue of aircraft.

During the 1980-1990s with many aging aircraft replaced with new aircraft (and perhaps also due to the extremely complex formulation of the problems), RAAF support for risk methods applied to airframe fatigue was not as strong, and the knowledge base was reduced within DSTO as people retired. However since the year 2000, with many more aging aircraft, the RAAF are once again interested in risk assessment methods to assist decision-making in support of these aircraft, in conjunction with the semi-probabilistic methods specified in the aircraft standards.

Thus in 2000, use of probabilistic methods was renewed in DSTO for initial analysis of the multi-site non-interacting fatigue of the lower wing stringers of the aging B707 aircraft [18][19], using the USAF program, PROF. This work showed that, while PROF was adequate for transport aircraft with small extrapolation, it did not really cover all the variables of interest and only approximately calculated the probability of failure.

2.1 Computer Programs for Calculating Risk

Early on, reliability programs were fairly complex, with a strong emphasis on reliability but less on fatigue. Fatigue in those days was not as well understood. So these programs had mixed methods of failure such as fracture and time out. Time out failure mode was used to fail the aircraft in the analysis when the fatigue life was exhausted, which is a duplication of the fracture method of failing. Now we understand fatigue is predominantly about crack growth from an initial defect to final failure, in some ways the analysis becomes much simpler. This relative simplicity, combined with advances in general purpose data analysis packages such as Matlab and S-Plus/R, means that specific reliability programs are no longer required. In addition, based on DSTO's specialist areas in fractography and fatigue testing, fractographic analysis of cracks from fatigue tests enables estimates of crack growth curves and initial flaw sizes to be more precisely obtained, underpinning the renewed interest in probabilistic assessment in which the contributions of all these areas are combined.

During the 1970s ARL had a large number of people involved in the development and application of risk methods to aircraft fatigue. It culminated in a number of risk analysis programs of which the best known is NERF (Numerical Evaluation of Reliability Functions) [20]. This program solved the numerical integration of the risk function to determine the probability of failure from a number of different failure criteria, although less effort was deployed on methods of acquiring useful data to support application to real cases.

Since NERF is no longer used we will limit our discussion to only consider PROF and SAIFE.

2.1.1 The USAF program PROF

In the 1980s the USAF developed a program PROF (Probability of Fracture) for solving the risk of structural failure [21][22][23]. Whilst this solves for the risk of failure, it does not strictly point out to the user how the risk of failure of a single item differs from the risk of failure of the complete aircraft. Here is a summary of a user's perspective of how to use PROF.

How do I start running PROFV2 ?

PROF is executed in Microsoft Windows. The PROF program can be found under the file name **PROFv2 3-24-99.EXE**. Copy this file into the directory on your hard drive and execute it. This will give you the PROF executable. PROF interfaces to an excel spreadsheet which contains all the input and output information, therefore when PROF is started up it automatically opens an excel spreadsheet as well. Using the PROF interface, all the required data files can be read in for analysis.

What input does PROF need ?

You need to have the following information to run PROF: The following key inputs are required for PROF:

Probability Distribution Functions

1. parameters for a probability distribution of the initial flaw sizes or the flaw sizes at a particular time (this must be for one of the distributions that is used by PROF, such as the normal, lognormal, Gumbel),
2. parameters for the normal distribution of the fracture toughness,
3. parameters for the log-odds (also known as the log-logistic distribution) probability distribution that describes the probability of detection. The log-odds probability distribution is another standard distribution that is useful in representing probability of detection.
4. Maximum Stress which is modeled with a Gumbel probability distribution which fit the largest expected load in a flight,
5. (Gumbel)

Deterministic Data

1. Normalised stress intensity function

2. master crack growth curve,

Aircraft Related Data

1. Number of locations per aircraft
2. Number of aircraft
3. Hours per flight
4. Specific times at which inspections occur, and

What do I get out of PROF ?

1. A plot or table of instantaneous risk versus flight time.
2. A plot or table of the cumulative probability distribution of the crack sizes that will fail. This is useful in determining how good the inspection method has to be. A large spread of failures means even small cracks can fail, thus requiring a very good inspection technique to lower the risk.
3. The flaw size distribution at each inspection time interval. This should be compared with the actual flaws found as this is the only means of reliably checking the risk calculation.
4. The latest version of the program (Version 2) does not output the total or interval risk despite what the manual says.

What can go wrong with running PROF ?

1. The risk can be very sensitive to the tails of the distributions. Because PROF uses tabular inputs this means the tails may not be well defined. Best to fit a proper distribution and obtain values from that to pad out the tail when inputting the data to PROF.
2. The probability of detection curves (POD), cannot be input into PROF using a tabular input, instead PROF uses its own in-built functions to approximate the curves. It is therefore important to make sure that the POD curve has a good fit at the large crack sizes.
3. The integration method in PROF does not seem very accurate when risk of failure is smaller than a certain level (typically this is about 10^{-10}). This is most likely due to fixed integration parameters being used for the numerical integration. Thus, if the risk levels are smaller than this very early on in the life of the aircraft there may be discontinuities in the shape of the curve that are an artifact of the solution accuracy.

The equivalent flaw size is obtained from teardown of the aircraft, and then fracture mechanics is used to translate the cracks back to zero time to obtain the EIFS, or to a common flight hour that minimises translation to obtain the EIFS distribution. The important point is that the crack size information must be translated to a common flight hour to be used in the risk analysis. The Aircraft Structural Integrity Managment Plan (ASIMP) managers should collect fleet crack information data such as airframe location, crack size, flight hours, and equivalent flight hours based on actual usage.

Berens, Hovey and Loomis [23] showed the EIFS distributions obtained from a number of different aircraft such as F4, F15 and A-10, where there are significant differences between each of the distributions. They stress that the tails of the distributions have a significant impact on the risk analysis calculations.

The SFPOF is the instantaneous risk at some time in the aircraft life and is frequently referred to as the hazard rate.

The USAF practice for risk threshold has historically been a SFPOF of 1×10^{-7} per flight, adequate for long term operations. Exposure should be limited when the risk is above this level (when the risk level is between 1×10^{-7} and 1×10^{-5}). A risk level above 1×10^{-5} should be considered unacceptable.

The PROF [21] program is the tool used for USAF risk assessments. It has previously been noted that the formulation for the SFPOF in PROF is not strictly the hazard rate. It is based on growing a population of cracks forward to the time of interest and then determining the probability of this population of cracks failing in the next flight. The formulations given for the SFPOF, however, while they are correct, are at odds with the formulation given in the PROF manual [23].

Thus the program assumes there have been no failures to date even though very large cracks may exist and would certainly have failed in earlier flights. Thus the term unfailed SFPOF was coined to describe the probabilities calculated by the PROF program [21][22]. While at low time the unfailed SFPOF is very close to the true SFPOF, at high times the unfailed SFPOF approaches the cumulative probability of failure (since as the cracks grow very large, it is almost certain that all of them will fail in the next flight). Thus the PROF calculation transitions from the true SFPOF to the TPOF and is thus overly conservative. In addition, the way numerous failure locations are incorporated into the risk analysis is generally incorrect. It assumes all the risk of individual locations is combined together.

It is also given that *to calculate the single flight probability of a fracture from any one of the k equivalent elements (stress raisers) in a single airframe at T flight hours* $\mathrm{POF}_A(T)$, *it is assumed that the fracture probability between elements are independent* (as noted elsewhere this is usually not the case):

$$\begin{aligned} \mathrm{POF}_A(T) &= 1 - [1 - \mathrm{POF}_E(T)]^k &(1)\\ &\approx k\mathrm{POF}_E(T) &(2) \end{aligned}$$

The last approximation is only valid when $\mathrm{POF}_E(T) << 1$.

However calculation of the combined probability of failure is only true if each of the locations are independent. Since the risk of failure is a combination of having a crack big enough to fail with a certain load, the loads for each location may not be independent. For example, all the critical locations in a wing will experience the peak load at the same time. Thus, assuming they are independent is again overly conservative.

The SFPOF at a single stress raiser in a single flight at T hours is given by:

$$\begin{aligned} POF_E(T) &= \Pr[\sigma_{\max} > \sigma_{cr}(a, K_c)] &(3)\\ &= \int_0^\infty \int_0^\infty f_T(a)\, g(K_c)\, \bar{H}(\sigma_{cr}(a, K_c))\, dK_c\, da &(4) \end{aligned}$$

where

$f_T(a) =$	probability density function of crack sizes at T flight hours
$g(K_c) =$	probability density function of the fracture toughness of the material
$\bar{H}(\sigma_{cr}(a, K_c)) =$	$\Pr[\sigma_{\max} > K_c/(\beta(a)\sqrt{\pi a})$

2.1.2 The FAA program SAIFE

Prior to the development of PROF, the Federal Aviation Administration (FAA) also developed their own program for risk assessment, with a focus on wide body aircraft which were then appearing. The FAA in conjunction with airframe manufacturers and air carriers were trying to improve the structural integrity and inspection efficiency of civilian transport aircraft. To do this the FAA developed a computer program called SAIFE - Structural Area Inspection Frequency Evaluation [24]. The aim of this program was to quantify the process they currently went through to set inspection intervals for a new aircraft type. It was also hoped the process would be more rigorous so that there was less subjective assessment.

The SAIFE program accounts for the following factors:

1. the aircraft design analysis,
2. aircraft full scale fatigue testing,
3. production, service, and corrosion defects,
4. probability of crack or corrosion detection, and
5. aircraft modification economics.

The program allowed for these effects through making a simulation of the process based on the Monte Carlo approach for incorporating the effect of scatter. The program created a scenario, where each of the factors comes into play. By repeatedly running through the scenario, a safe and economical approach to operating the aircraft can be achieved. So once a simulation is set up, different inspection programs can be trialled to evaluate their effectiveness.

In addition to the development of the computer program, there was also research conducted into the inputs, such as the effectiveness of inspections.

Volume III of [24] contains inspection input, survey data, and Mechanical and Reliability Report (MRR) data. Similarly, Volume V, Results of Model Demonstration, presents the results of the program applied to a hypothetical aircraft but essentially similar to a B-747 which was just coming into service at the time the program was developed. It then compares these results with the service experience of operational aircraft. It does not compute actual risk levels for the user (although in theory it could), but provides the expected number and severity of cracks that can be expected during the life of the aircraft.

The primary objective of the SAIFE project was to develop a program for evaluating inspection intervals. The program is tailored to be designed for wide-body aircraft such as the B-747, but it may also be run on individual components, such as for a critical area on a wing. For each element simulated, SAIFE generates the number of defects due to cracks, corrosion, production damage and service that occurred during the service life of the aircraft.

The simulation output includes the number of cracked and corroded areas detected. This can be used as a comparison between actual levels detected and the expected number to see if they agree.

It was hoped that SAIFE could determine the inspection intervals on a more objective basis than was typically used at that time. The project assembled all the logic used to establish inspection frequencies into a single simulation program, capable of investigating the interactions between the primary aircraft service life factors - ultimate strength, fatigue life, flight loads, production and service damage, corrosion, probability of defect detection and modification economics.

SAIFE uses a series of probabilistic distributions and deterministic equations to simulate the logic sequence that considers all the subjective elements currently considered in arriving at an inspection program. It is based on the logic developed by Aderjaska [25].

The input data consists of three categories of information:

1. fleet information
2. aircraft design data
3. historical data

Fleet information identifies type of aircraft e.g. B747, the number of aircraft in the fleet and the expected service life of the aircraft. The aircraft design data includes a complete breakdown of the aircraft structure into components such as a wing structure or a fuselage frame. For such a breakdown the fatigue life, ultimate strength, crack growth rate, corrosion resistance and corrosion growth rate for each element must be determined.

The historical information is particularly relevant for obtaining corrosion data and comes from the MRR. After 1972 the MRR became known as the Service Difficulty Report (SDR). All United States carriers are required to submit an MRR/SDR whenever they repair a defect. Volume III presents the MRR data covering the ten year period of 1963 through to 1973 which was analysed for inclusion into the SAIFE program.

To relate the design lives to the actual fatigue lives, they use results of a report by Raithby [26]. This report finds that in about ten percent of designs the actual life of the type is less than the predicted life. However for the analysis used in SAIFE, they assume on average the predicted life equals the actual life. Thus in fifty percent of designs the service life is overestimated. This may be excessively conservative.

For each element the Monte-Carlo method is used to select a predicted life multiplier. The actual fatigue life is then determined from the equation

$$\text{actual life} = \text{predicted life} \times R \tag{5}$$

This is very similar to the approach used in the preliminary F/A-18 centre barrel risk analysis [14][27], where the accuracy of the predicted fatigue life based on the individual aircraft fatigue tracking was taken into account.

The SAIFE logic uses the two parameter Weibull distribution as developed by Freudenthal [28] to allow for the variability in fatigue lives. The two parameter Weibull was chosen because it was more flexible, and could easily be modified to take into account any changes in material technology, even though it was not as accurate as some other distributions in representing the tails of the distribution compared to actual data.

In commercial jets the average fatigue life is taken as twice the safe fatigue life. This is the criterion used on the DC-10 and the B-747. The Weibull distribution is used to model the variation in fatigue life by introducing a fatigue life factor.

The SAIFE program simulates the fatigue test result by assuming the fatigue test failure takes place after a test period equal to the estimated average fatigue life divided by the fatigue test acceleration factor. If the estimated average fatigue life is less than twice the service life when a fatigue test failure occurs, a production modification is developed. If the estimated average fatigue life is less than the service life a retrofit modification is developed for those aircraft already in service. The inspection frequency is increased from the time the element reaches 80% of its estimated average fatigue life.

If a modification is tested, it is assumed that the demonstrated life will be equal to the predicted average life. If a modification is not tested, then it is assumed that the estimated average life will be subject to the same uncertainties as for the original design.

Service data are used to determine the rate of occurrence of service damage, production defects and corrosion so that it is accounted for in the simulation. From this one can define the rate at which production defects occur, the rate at which service damage occurs, the rate at which corrosion occurs and grows, and in turn, the effects of production defects and service damage on fatigue life.

The loads in SAIFE are based on a sample of 2000 hours of operational usage that included gust and manoeuvre loads on the wing and fuselage for a typical aircraft obtained from NASA flight trials. Civil transport aircraft have four levels of inspection, each with a different probability of detection. The inspection regime ranges from the pre-flight inspection which has a very low probability of detection, through the service phase inspections, to the overhaul which has a much higher probability of detection.

A sensitivity study is performed to determine the effect of the variables to optimize the inspection program. The SAIFE authors maintain that while there are many variables through which the structural safety can be gauged, the structural element failure rate is the most revealing.

A detailed run of the program was provided in Ref. [29]. This was for an example aircraft based on data which was taken from a combination of the B747 and the DC-10 wide-body aircraft. The simulation analysed over 1300 locations on the aircraft to determine the rate of cracking that would be detected at each inspection level. This was then compared with typical data obtained for the previous two aircraft. Initially the SAIFE program predicted too many cracks would be found. But after some fine tuning of the probability of detection, and allowing for cracks that originate in a hidden location to be seen once they were sufficiently big, the program produced results that were in broad agreement with actual data. One drawback however, was that the program only printed out crack locations and numbers of cracks to be found at each of the different inspection levels. It did not indicate the severity of the cracking or whether the more severe cracks would result in the loss of the aircraft. The aim of the program it seems was to optimize the inspection levels. However this does not seem to be optimized with the view to maximizing air safety. It may have done this implicitly, but the explicit level of risk and how this changed as a function of time was not given in the sample output data.

2.2 Safe life and damage tolerance

Around 1970 the USAF set about changing the way the structural integrity of its aircraft was managed [30]. This was brought about by a number of problems in new and existing aircraft which led to an attempt to manage the aircraft from "the cradle to the grave". Part of this change was a revision to the airworthiness standard MIL-A-8860 which influenced the way the aircraft was designed. The use of the standard in conjunction with tight schedules, which limited testing, resulted in low design weights being necessary for components which in turn forced the selection of high-strength fracture sensitive materials and use of high design stress levels. Studies at the time confirmed that most structural failures were as a result of fatigue or corrosion induced fatigue, initiating from pre-existing material and fabrication deficiencies.

The emphasis in MIL-A-8860 was on initial static strength and a safe fatigue life, using scatter factors to account for variations in environment, material properties and initial quality. A scatter factor of four was used on the test life, similarly the times at which modifications were introduced were also based on a scatter factor of four. Because of failures in some aircraft (notably the F-111), the safe-life approach was abandoned and it became a requirement for a structure to survive with flaws from manufacturing or service. These requirements are specified in MIL-A-83444 for Damage Tolerance Design and MIL-8866B for Durability Design. Testing to ensure compliance is given in MIL-A-8867B. However, although standard initial defect sizes were then used for damage tolerance analysis, the original flaw in the F-111 still far exceeded the flaw size used to ensure damage tolerance, so even this change would not have prevented the original F-111 failure. Instead, F-111 life management relied on an inspection-based approach.

The fatigue test guidelines were formed such that the most important purpose of the full-scale fatigue test is the identification of critical fatigue areas. In addition the fatigue test can be used to verify the initial flaw quality/distribution. The fatigue test also demonstrates the durability of the airframe to fulfil its service life.

3 Calculating the probability of failure

3.1 Fundamental concepts and probability distributions

Early methods of probabilistic analysis [31] for aircraft focused on static strength considerations. Failure was then relatively easy to determine by determining the probability distribution of the strength of the structure due to variation in the input variables and finding the probability that the strength is less than the applied loading. Interestingly, it was recognised even at this stage that each component could have a probability of failure and that the total probability of failure of the complete aircraft was a combination of the individual probabilities. For design purposes it was suggested that the reliability should be distributed among the components of a structure so that the probability of failure of each component is proportional to its weight. A safety factor can then be defined in probability terms as the ratio of the probable strength to the probable maximum force. If the applied and failing loads are accurately defined, then a safety factor of 1.01 would give high reliability, whereas if there is large variability, a much higher factor would be required.

However, in a recent risk assessment of critical cast titanium components on the F-22 aircraft [32], it was assumed that all the risk of failure was consumed by the cast components. It shows the difficulty even now, of allocating acceptable risk levels to individual components.

The most common essential elements of a structural reliability analysis system based on fatigue failures are [33]:

1. a crack initiation and growth model,
2. a residual strength model to calculate the strength of the structure with cracks of arbitrary length,
3. a structural inspection and crack detection model, and
4. a maximum flight loading model which calculates the probability of having a specified maximum flight load.

In addition to the usual measures of probability of failure such as the cumulative and the single flight probability of failure, Freudenthal [8] has proposed another way based on the probability of time to first failure in the fleet.

This method is based on extreme value statistics, but uses the minimum extreme value distribution. A feature of this is that the distribution can be determined based on physical requirements.

This approach is certainly valid if the failure distribution is regular and if there are sufficient failures to enable the characterisation of the distribution of the first to fail in a group of a certain size. He goes on to show how generally one test, such as a full scale fatigue test, is sufficient to characterise this distribution. However, this is based on the assumption that the parameters to characterise the base distribution are known exactly. The examples he gives are based on parameters derived from component tests. This then assumes that all components are identical and all failures come from the same population. This may not always be the case for an aircraft.

In addition, because the extreme value distribution is focused on the centre of the distribution instead of the tails it is somewhat insensitive to the distribution used. However, extreme value theory indicates there should be only three extreme value distributions (Gumbel, Weibull, Frechet). Freudenthal puts forward that fatigue failure, because of its very nature will always take a definite number of cycles to first failure, so that it must be the Weibull distribution (the other two extreme value distributions do not have a finite lower bound). The Weibull distribution is also known as the Fisher-Tipper Type III asymptotic distribution of the smallest extreme. The Gumbel is referred to as the Type I extreme value distribution and the Frechet is the Type II.

The distributions of the smallest and largest values in a sample will in general be functions of both the sample size and the nature of the initial distribution, Provan [34]. However, if the number of samples becomes large, and the tail of the initial distribution is of the exponential type (which includes the exponential, the normal and lognormal distributions), Cramer in 1964 showed that the smallest or largest value converges asymptotically to the so-called Type I extreme value. Since this distribution was extensively used by Gumbel in his study of extremal phenomena, it is also known as the Gumbel distribution.

The Gumbel distribution of the *smallest extreme* is characterised by the distribution tending to zero as the random variate tends to minus infinity. The Probability Density Function (PDF) is given by:

$$f(t) = \frac{1}{\delta} \exp\left[\frac{1}{\delta}(t-\alpha) - \exp\left(\frac{t-\alpha}{\delta}\right)\right] \tag{6}$$

where $-\infty < t < \infty$; $\delta > 0$; $-\infty < \alpha < \infty$

$$F(t) = 1 - \exp\left[-\exp\left(\frac{t-\alpha}{\delta}\right)\right] \tag{7}$$

The Gumbel distribution of the *largest extreme* is such that the distribution tends to zero exponentially as the random variate tends to plus infinity. Which is given by:

$$f(t) = \frac{1}{\delta}\left[-\frac{1}{\delta}(t-\alpha) - \exp\left(\frac{t-\alpha}{\delta}\right)\right] \tag{8}$$

$$F(t) = \exp\left\{-\exp\left[-\left(\frac{t-\alpha}{\delta}\right)\right]\right\} \tag{9}$$

How large the sample size N is before the maximal distribution asymptotes to an extreme value distribution, depends on the type of the initial distribution. Fewer observations are required for the maximal distribution to approach the Gumbel distribution if the initial distribution is an exponential rather than a normal.

The exponential distribution can also be considered:

$$f(t) = \frac{1}{\delta} e^{-t/\delta} \tag{10}$$

and the cumulative distribution is given by:

$$F(t) = 1 - e^{-t/\delta} \tag{11}$$

where $1/\delta$ is the hazard rate. The hazard rate is defined as the limit of the failure rate as the time interval approaches zero. Thus the hazard h is defined as:

$$h(t) = \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t \, [1 - F(t)]} = \frac{1}{1 - F(t)} \frac{dF(t)}{dt} = \frac{f(t)}{1 - F(t)} \tag{12}$$

The above distribution is characterised by the property known as a complete lack of memory. In other words, in the case of an exponentially distributed time to failure, the hazard is constant and has no memory and depends only on the specific time interval and not on the previous history, e.g. whether the component has survived 100 or 1000 hours. The exponential is a special case of the Weibull and Gamma distributions. The Gamma distribution is similar but generalises the exponential so that it represents the reliability of a system with a number of components, the failure of any one of which will lead to the failure of the system at a constant failure rate. Like the exponential, it also has no memory.

The normal probability distribution is also popular, however, its use as a reliability (or time to failure) distribution is questionable since it includes negative values in the time to failure. Lognormal is particularly of interest if the failure is due to the growth of a fatigue crack [34]. If the rate of crack growth is proportional to the crack size then the time to failure will be lognormal in the limiting case.

3.2 Methods of determining probabilities of failure

There are two sorts of uncertainties in a risk assessment [1]: Aleatory or irreducible risks, and epistemic or reducible risks. Aleatory risks consist of elements in the risk assessment which cannot be improved upon, including loads, geometry, and material properties and defects. Epistemic risks consist of uncertainties which can be improved upon with further work, which include, analysis models, strength models, statistical data bias and human error. Some elements of fatigue and risk assessment have both epistemic and aleatory components. An example is crack growth, which has behavioral trends amenable to deterministic modelling, but also has variability about the trends, which is apparently irreducible. The variability is normally dealt with by allowing the parameters of the model to be variables.

This report notes that while probabilistic models can be created given sufficient data, where there are two or more random variables, it is important to establish whether these variables are statistically independent. In other words the degree of correlation between the variables needs to be determined.

Orisamolu [1] feels that the requirement for establishing independence and allowing for it in the calculations is not accorded due importance by the analyst. In fact, most of the reliability software does not allow for correlation between the variables. Because there is generally insufficient data to adequately perform a reliability analysis, Orisamolu recommends other techniques also be used. These include confidence bounds, dual uncertainties and the Baysian approach.

The USAF developed the probabilistic approach between about 1970–1990 [35][36]. Because the Aircraft Structural Integrity Program (ASIP) focuses on the durability and damage tolerance requirements due to fatigue damage and crack growth they conclude that probabilistic fracture mechanics is a cornerstone of any probabilistic risk assessment and follows the approach of Yang [37] for assessing the durability and damage tolerance of an aircraft. Orisamolu suggests the Canadian Forces use for future risk assessments the model used at the time by Canadair. This is the lognormal random variable model, which lumps all the uncertainty in the crack growth into a single random variable that is assumed to be lognormally distributed:

$$\frac{da(t)}{dt} = X(t)\, g(\Delta K, R, a) \tag{13}$$

where $X(t)$ is a random variable that accounts for the variability in the crack growth accumulation. However, Orisamolu also felt that more flexibility was required than provided by this simple model alone.

In addition, because of the uncertainties of the inputs to the reliability model, it is also important to determine the sensitivities of the individual parameters to the final result. System reliability methods also need to be developed to deal with the physical reality that there may be multiple flaws and defects in a given structural detail. The problem of multiple hot spots can be dealt with by simplified procedures that neglect the effects of possible interaction. e.g. load shedding and subsequent stress redistribution to adjacent members, or statistical correlation between the parameters that drive the fatigue failure at these locations. However, this approach may give overly conservative results.

Generally there are two approaches to determining the probability of failure. These are analytical or numerical. The analytical approach generally involves solving the integration problem that describes the probability of failure of the system and so adds up to find the total probability of failure at a given time. Because the solution of these problems can be quite complex there are simplifications that are made to allow the problem to be solved numerically. These include FORM/SORM which are the first order and second order reliability methods. Essentially these approximate the exact distributions of failure by equivalent normal distributions. The combination of many normal distributions is particularly easy to solve.

However the most popular way is to perform a Monte Carlo simulation. This involves repeatedly performing a deterministic analysis starting from initial conditions selected at random from the distribution appropriate to the input variable. This is repeated until a sufficiently large random sample of failed and unfailed items has been calculated so that the probability of failure can be determined directly from the sample proportions. According to Hammersly [38], this technique was first devised by Ulam and Von Neumann during World War II. Monte Carlo methods were originally developed for use on the atomic bomb to solve problems in random neutron diffusion in fissile material. The name originates from the Monte Carlo casino which Stanislaw Ulam's uncle used to frequent.

Hence when a deterministic program for evaluating failure exists, such as a crack growth or finite element program, it can be incorporated into a larger probabilistic program. This is most easily done if the interface to the program is via a command line or can be called from another subroutine. This allows a probabilistic program to call the specific deterministic program with the appropriate inputs as required. However, if the sole means of running a deterministic program is via a graphical interface which requires direct human input, then it cannot be used. Thus a code written for future use should have in mind three levels of input - subroutine, command line and graphical user interface.

3.3 NASA Recommendation

NASA have developed a series of brief guidelines [39] for the use of probabilistic fracture mechanics in spacecraft such as the Space Shuttle. They suggest a simple technique which uses the Paris or modified Forman equation for crack growth.

NASA define a limit state function M as:

$$M = \int_{a_o}^{a_c} \frac{da}{Y(a)^m(\sqrt{\pi a})^m} - C\nu_0(T - T_0)E[S^m] \tag{14}$$

where

a	= crack length
a_0	= initial crack size ($a_{t=0}$)
a_c	= critical crack size at which fracture occurs
$\nu_0 T$	= number of stress cycles in time T at average frequency
$\nu_0 T_0$	= time to crack initiation
C, m	= material constants in the Paris equation
$Y(a)$	= geometry function for physical problem under consideration
S	= far-field stress range
$E[]$	= expected value (mathematical expectation)

Failure is defined to occur when the critical crack length a_c is exceeded, so that at failure $M \leq 0$. The probability of failure is then the probability that the limit state function is equal to or less than zero, which can be expressed as:

$$P_f = P[M \leq 0] \tag{15}$$

This integral must be evaluated numerically in all but the simplest of cases. It should be noted that all terms in Equation (14) can be treated as random or uncertain. Subsequent sensitivity analyses can be used to determine which variables may be considered as fixed (deterministic).

4 Initial Defects

It is difficult to accurately predict the rate of crack growth with traditional models, down to sizes that approach the initial defect. Most crack growth programs acknowledge a lower limit of crack growth of approximately 1mm. Below this the predicted rate of crack growth is much lower than that observed. Thus it is unreliable to use conventional crack growth programs to predict the crack growth from the initial flaw size, unless there has been some specific modification to allow for what is known as the small crack effect.

For aluminium type materials, initial flaws may be intrinsic material effects such as porosity and inclusions of the order of 1–50 μm (1 μm or micron, being 1 thousandth of a millimeter), or extrinsic due to drilling, etching or corrosion, which are typically larger, and of the size of 50–200μm. For determination of the probability of failure this does not present such a problem. All that is required is that there are sufficient data to characterise the initial flaws such that the predicted distribution of final crack sizes produced is statistically correct. In fact, the only time the predicted crack size needs to be strictly correct is when it interacts with other things in the real world such as the ability to be detected or is of such a size as to potentially cause failure. This approach of modelling the entire crack growth process is known as the EIFS and was developed by Rudd and Gray [9][10][11]. An alternative to using an initial defect size is to use a time to a finite crack size instead.

The development of the understanding and modelling of fatigue of materials and its variability has progressed over many years from a macroscopic/phenomenological approach to a microscopic/mechanistic approach. Initially, fatigue was modelled as simply a total failure process, and the path from the start of life to final failure was not considered. Statistics and models were developed on the statistical distribution of the time to failure and how it is affected by load amplitude, load sequence, material, etc. These whole-of-life statistics and models underpin the safe life approach to the management of aircraft safety.

However, it has been well known from the start of fatigue research that fatigue is a process in which a crack initiates at a particular location and grows slowly and progressively under service loading to a size at which the residual strength of the structure is insufficient to sustain the prevailing loads, and rapid failure occurs. The ability to model the key stages of this process (initiation, growth and fast fracture) has been the subject of much research. This modelling underpins the damage tolerance approach to the management of aircraft safety. The fast fracture stage was the initial domain of fracture mechanics research, and may be regarded as being at a mature state of understanding and modelling, although the progression from the onset of fast fracture to total structural failure often involves many assumptions.

In recent years, the main focus of the research has been on modelling the crack growth stage. There has been less recent research on the initiation stage. The EIFS approach effectively sweeps up all the complexity in the initiation process of a crack being triggered by some defect and evolving into a regular shape for self-similar growth. EIFS is therefore a construct, a projection of self-similar crack growth back to the start of loading, and is only loosely correlated with the defect size at which the crack began. This limits the portability of EIFS across materials, structural configurations, load spectra and even stress amplitudes in some instances. Nonetheless, the EIFS construct is a useful, practical engineering approach to dealing with the initiation stage, pending the outcome of more research. EIFS contains both aleatory and epistemic contributions, but is treated as a wholly aleatory entity. Future research should separate the contributors to EIFS and appropriately model them. This remains as perhaps the final frontier of fatigue research, leading to the ability to do accurate predictions of the fatigue of structures.

Manning, Yang and Rudd [40] give examples of how to calculate EIFS using fractographic data combined with a number of different growth models. They show how to determine initial flaw distributions from what they hope would become a standardised coupon test. They proposed coupon tests with no load transfer (through a rivet) and 15% load transfer to determine the initial flaw quality for fastener holes. A fastener is required in a hole to fully restrain the sides of the hole. The amount of restraint the fastener provides depends on the type of fastener and how well it fits the hole and the stress level it is subjected to. However, they were unsure whether open hole specimens could be used conservatively to determine the EIFS of fastener holes. No hard and fast rules were given for the number of specimens to be tested, other than saying it depends on the type of specimen to be used, such as whether it is a single hole or a multiple hole and the design variables which need to be accounted for, such as material, fastener type and fit, fastener diameter, bolt torque, manufacturing variations, fretting environment, stress level, load spectra etc. Based on this alone it would seem the number of coupons required to determine all these variations is unrealistically high.

Manning also recommends that crack growth data should be obtained from fractographic examination of the cracks. However, he cautions that the minimum crack size that needs to be determined from fractography should be consistent with the minimum crack size for use in linear elastic fracture mechanics and is typically of the order of 0.127 mm. He suggests that a crack in the region of 0.25-1.25 mm would be reasonable to assess the initial flaw quality of fastener holes.

In deriving the optimum initial flaw distribution, the best combination of parameters should minimize the error between predictions based on the candidate EIFS and the actual fractographic results. Ideally the EIFS distribution should be fitted so that the best fit to the data is at the upper tail portion of the population, because it is the largest EIFS that will govern the durability of the structure.

In fitting the crack growth data, Manning, Yang and Rudd [40] were happy to use a variety of linear and log-linear approaches. They solved for the parameters of best fit using least squares. It is necessary to determine a master service crack growth curve for each region that is to be analyzed. For a reliable prediction of the crack exceedance distribution, the crack growth master curve should be compatible with the EIFS distribution used. Guidelines for this are available in Manning [41] [42]. The EIFS should then be grown forwards in the same way the fractographic data was extrapolated back. This requires that the master crack growth curve be based on the same law that was used for the normalised EIFS master curve.

The article by Manning, Yang and Rudd [40] gives an example of a risk analysis of the lower wing skins of a fighter aircraft. This was a real aircraft that was durability tested to 16 000 flight hours which was equivalent to two service lives, using a 500 hour representative block spectrum. Both wings had the same loading and at the end of the testing a teardown inspection was performed inspecting all of the holes in the lower wing skin using eddy current. Those holes that were found to have an indication were subject to fractographic examination. They found 26 holes cracked in the right hand wing and 7 holes cracked in the left hand wing that had cracks greater than 0.76 mm at 16 000 hours.

The initial flaw-size distribution was based on a three parameter Weibull distribution. They also pooled this data with coupon test results by testing a series of coupons at three different stress levels. For the risk analysis they divided the lower wing skins into 10 stressed regions, and determined a master crack growth rate for each region. Once they had the crack growth rate curves and the initial flaw size distribution they were able to determine the fastener hole crack size exceedance distribution for the wing at 16 000 hours. However that was the extent of the analysis. There was no attempt at defining probability levels of failure. While they were aware that the largest cracks are the critical ones they do not consider that only the largest crack is really of significance.

Although there are a number of publications determining the distribution of initial flaws for typical holes, the determination of initial flaws has largely been neglected [43].

However, there have been six major investigations undertaken to determine the distribution of EIFS or EPS for a number of materials. Refs. [44][45] were based on teardown data of aircraft fastener holes. One investigation [46] was based on defects due to the etching process used during the manufacture of F/A-18 bulkheads. Another substantial investigation by Molent et al. [47] developed initial flaw sizes for the materials and processes used in the manufacture of the F/A-18 bulkheads.

One drawback with teardown data collected from aircraft is that the loading sequence is generally unknown and hence it becomes difficult to determine the crack size as a function of time. Conversely, by using coupon tests with a repeated fatigue load sequence and fractography, it is possible to determine the crack size from very small initial flaws (microns) until failure (millimeters), but at the expense of moving away from “representative structure”.

For fighter aircraft where the majority of the crack growth consists of growth in the sub-millimeter range it is important to characterise the distribution of initial flaws accurately in these regions. However, nearly all attempts to quantify EIFS have focused on flaws originating from fastener holes.

Wang [44] conducted a test on 2 coupons made from 2024-T3 aluminium alloy with 24 unfilled holes in each coupon. By cutting out the cracks when small with a hole saw, he was able to continue testing the panel until all holes had been removed and examined (an interference fit plug was installed in the excised hole to prevent further cracking). In addition, by testing with a sequence containing identifiable marker loads it was possible to determine the rate of crack growth. However because of the time consuming process of performing fractography with an SEM he was not able to determine the crack growth curves for each hole and had to resort to using an average crack growth curve to determine the EIFS. If the marker band spacing was greater than a micron, long working distance optical microscopes would have provided a quicker method for measuring the rate of crack growth.

As is commonly noted for small cracks the growth was found to be exponential [46][48]. Also noted was that the growth rate of a crack was influenced by cracking on the opposite side of the hole. The scatter in the final fatigue life was able to be described in terms of

EIFS, micro-crack growth, and macro-crack growth. The origins of failures were found to be from tool-marks, burrs and inclusions.

Kiddle [49] provides crack nucleation count results from 134 fatigue tests on lug specimens of four different aluminium alloys.

4.1 Types of structural details

There are three different types of structural details that are fatigue critical. These details are different because of the way they affect the number and position of initiation sites, how cracks from these sites link up and grow, and the NDI means used to detect the presence of these cracks.

The three main types of features which commonly initiate cracks because of the higher stress concentration they cause are

1. holes
2. fillets
3. lugs

It is debatable whether lugs should be treated any differently from holes. Lockheed [50] conducted a comprehensive survey of fatigue cracking from lugs, and after questioning many NDI inspectors, they were not uniformly of the opinion that lugs were simply another form of hole. Of note is the fact that there are only a few lug holes so they may be expected to receive better NDI treatment. In addition the degree of hole clearance and diameter to thickness ratios have a significant effect on the number and position of the initial flaw sites that result in fatigue cracks. Finally, lug joints usually involve high load transfer through the pin, whereas other joints with holes usually involve clamp-up action with faying surface friction load transfer and often there is a high bypass stress compared to the stress due to load transfer. Of course, there are also many open holes in an aircraft structure.

In another extensive review of failures of lugs in service, Kiddle [49] found when one side of the lug hole was cracked the other side was also cracked. (All but 2 of 155 specimens tested were cracked on both sides of the lugs). Kiddle also found there was a strong trend toward increasing the number of flaw origins with increasing stress. This may be caused by higher stress levels being able to crack inclusions present in the material, which then grow into fatigue cracks. At lower stress levels the inclusions remain intact. Although the finishing process of the material will also dictate how many inclusions are broken.

5 Crack growth

5.1 Probabilistic Effects in Crack Growth

Bent [51] looks at how to include random effects in a risk analysis of fatigue failure. A study by Yang [52] looks at crack propagation in fastener holes of aircraft structures and in a centre cracked panel with both subjected to random spectrum loading. Yang models the crack propagation rate as a lognormal process using:

$$\frac{dA}{dn} = C(t)\,(\Delta K)^m \tag{16}$$

where $C(t)$ is considered as a random process and m is a material property. In this study, Yang considered two extreme possibilities for the correlation length of the random process $C(t)$. For zero correlation length $C(t)$ becomes a white noise process, for an infinite correlation length it becomes a random variable. The white noise process introduced little scatter in the crack propagation time histories. Whereas the random variable option resulted in high scatter. The experimental results were found to be in between these extremes, hence the correlation parameter of the process was calibrated to reproduce the scatter observed in the experimental results.

Thus a simple lognormal random process model was able to reproduce the faster and slower rate of crack growth of individual crack length histories observed in experiments. Yang suggests that although good agreement was found by modelling the scatter with $C(t)$, which is variability of the crack propagation rate in time, if the scatter is due to material inhomogeneities, it would be more appropriate to model C as a random process as a function of crack length a. When time is introduced as the parameter, the results become dependent on the level of loading, since the crack propagates over more or less material within the same time depending on the loading level.

Models that consider C as a function of crack size have been proposed by Ortiz [53] where he uses:

$$\frac{dA}{dn} = \frac{C_1}{C_2(a)}\,(\Delta K)^m \tag{17}$$

where C_1 is a random variable describing fluctuations in mean behavior of different specimens, and $C_2(a)$ is a lognormal random process of crack size describing deviation from mean crack growth within a specimen. Ortiz found that the correlation length was close to zero and C_2 could be approximated as a white noise process. i.e. within a particular coupon the growth rate is uncorrelated. Additional variables can also be introduced into crack propagation laws to account for modelling errors [54].

5.2 Micromechanism of fatigue crack growth

The ability to be able to predict fatigue crack growth (whether deterministically or probabilistically) is hampered by a lack of understanding in how fatigue cracks grow. Although there exist hundreds of thousands of reports or articles on fatigue crack growth, much of this literature aims to document what happens and not why it happens. An understanding of why things happen is important because it sets the limits of applicability of various data sets, or at least includes variables that may need to be considered in the analysis. It also provides the starting point for a mechanistically based model. So we shall start with a brief summary of crack growth in aluminium alloys.

Cracks appear to grow from defects, very early on in the life of the component. When small, the method of crack advance appears to be by striations, such that the position of the crack front is indicated by the extent of a given striation [55]. (The proof of this fundamental statement is not strong so this may not be the case.) Striation advance seems to be caused at least in part, by an environmentally-assisted cleavage-like mechanism. The process is enabled by the presence of water vapor, which when in contact with freshly exposed aluminium (such as during the formation of a crack) will release free hydrogen atoms which are absorbed into the material allowing striations to be formed (either by increasing or decreasing the local ductility, the exact change is not known). The tension part of the load cycle causes the crack to propagate. There is a critical threshold of moisture/crack advance, such that below this level, striations (generally of types referred to as ductile or brittle) are not formed, and what appears is a wrinkled surface separated by irregular fissuring. This wrinkled surface has been termed as coarse striations, however the size and number of the striations do not correspond to the application of the individual loads. Crack growth below this threshold is an order of magnitude slower than above this threshold. Thus, in a vacuum environment such as found for internal flaws, which have no air or moisture, cracks typically have growth rates an order of magnitude lower than exposed external flaws.

Crack growth predictions [56] are generally based on relating the stress intensity at the tip of a crack to the stress intensity of cracks in specimens under constant amplitude loading. However to determine the equivalent stress intensity for a real aircraft is more complex and may be influenced by [43]:

1. crack geometry and location,
2. the types and combinations of external loads, for example tension and bending, and biaxial versus uniaxial loading,
3. fastener hole geometry (countersunk versus cylindrical),
4. the load distribution in pin-loaded holes,
5. fastener fit (neat fit, press fit, interference fit),
6. fastener hole residual stresses due to cold working and/or hole expansion, and
7. fretting in pin-loaded holes and at faying surfaces.

Most probabilistic risk assessments use linear elastic fracture mechanics with a simple law for predicting the rate of crack growth [33]. This may be the Paris law of growth (which is reasonable for constant amplitude growth typically found with transport aircraft fuselage pressurisation or ground-air-ground cycles and subject to the crack tip being away from stress concentration plastic zones).

However, when the spectrum is not constant amplitude, the effect of acceleration and retardation due to overloads and underloads can be important. However, whether it is worthwhile to consider retardation effects is also dependent on whether the exact sequence is available from the aircraft or whether a partial or typical sequence has to be used. If only a statistical description of the sequence is available then a number of predictions will have to be made to determine the effect of acceleration or retardation on the variability in fatigue life due to sequence effects.

All (to the author's knowledge) crack growth programs focus on deriving an answer which should agree with the average of a number of test results. No (standard) program seems to consider that knowledge of the accuracy of the answer is also important, particularly when used in a structural risk assessment. In these cases the accuracy of the answer can be included as just another uncertain variable. Therefore, knowing the accuracy of a fatigue life prediction becomes more important than predicting the average fatigue life since this variability can be incorporated in the risk analysis.

The issue of model inaccuracy can arise in a number of circumstances:

1. the model is too simplistic to capture the general crack growth behavior, e.g. an exponential crack growth model cannot capture the marked acceleration of crack growth rate at long crack lengths,

2. the model parameters are tuned to give a good fit to one load spectrum but the model is then inaccurate for other load spectra (in some respects this is a subset of the first case), and

3. a fixed model with clear deficiencies is required to be used, e.g. a fatigue tracking model that comes with an aircraft.

In summary, (as for the calibration of a measuring instrument) the systematic errors in a crack growth model need to be accounted for as well as the random errors. If the systematic errors are unpredictable, then it is possible to allow for them through an appropriate random variable.

5.3 Current state of predicting fatigue crack growth with a known sequence

The ability to be able to predict fatigue crack growth both deterministically and stochastically has made a number of important advances over the years[3]:

1939 Minimal Extreme Value Distribution developed

In 1939 Weibull [57] developed a statistical distribution determined experimentally from the strength of materials based on the weakest link. The distribution matched the failure strength of material. Later it was realised that this represented an extreme value distribution which related the maximum and minimums of a group of items. Weibull [15] went on in 1949 to apply this to represent the minimum time to failure for fatigued items.

[3]Note: this is only intended to be an overview. For a comprehensive review see Suresh [3]

1945 Miner proposed the cumulative linear damage model. This fatigue damage model which was originally proposed by Palmgren, and was re-discovered by Miner in 1945.

1951 First photographs showing striations

These were obtained by Zapffe and Worden [58] by using a normal microscope. Up until this point there was no documented evidence that people were using microscopes to look at the fracture surface. This was primarily because the irregular fracture surface required a large working distance lense. High magnification lenses generally have short working distances and so cannot be used to view most fractures. However, some fractures are very flat and could still be viewed with short working distance lenses available at the time. Modern lens have now been designed for longer working distances which makes viewing fatigue cracks easier.

1957 Stress Intensity Factor defined

Irwin in 1957 [59] characterised the state of the crack tip by the new parameter stress intensity factor. At the time, this was only used to predict the stress levels of static fracture. However the use of the stress intensity parameter was extended into fatigue crack growth assuming that the rate of crack growth was dependent on the stress intensity at the tip of a crack.

1958 One load cycle produces one fatigue striation

By examination of striation patterns produced in a laboratory environment from a number of specially designed loading sequences, Ryder [60] made the correlation that one load cycle produces one striation by looking at simple sequences with different load magnitudes. Later work by McMillan and Pelloux [61] showed that the striation advance was produced by the tension only part of the cycle. The compression load changed the appearance of the striations profile but did not affect the amount of advance.

1961 General Crack growth law by Paris and Erdogan

In 1961 Paris [62] proposed that the rate of crack growth for a particular cycle da/dn is related to the stress intensity and the stress ratio of the cycle. Paris and Erdogan [56] quantified this to the simple relationship that relates the change in stress intensity in the cycle to the crack growth rate intensity at the crack tip ΔK.

$$\frac{da}{dn} = C\Delta K^m \tag{18}$$

Where the material constants C and m have to be obtained experimentally. Interestingly, when first submitted for publication, Paris' paper was rejected by some of the top fracture journals of the day. This was probably due to a number of apparent problems with the formulation, amongst them that the units are inconsistent when the exponent m is not 2, and that the use of ΔK implies a perfectly plastic material (at least at the crack tip). This would seem to be because the memory effect in a crack is completely wiped out during the compression part of the cycle. Thus the start of the tension cycle is effectively starting from zero and giving rise to the crack advance being proportional to ΔK.

However despite these shortcomings, the Paris equation does seem to work over a large range of crack growth, albeit for largely constant amplitude cycling and large cracks.

1971 Crack closure proposed by Elber

Elber [63] proposed that the plastic zones in the wake of the crack tip actually hold the crack open, which has the effect of reducing the range of stress intensity and hence the driving force of the crack growth. Elber suggested a modification to the Paris law Equation (18) gives the effective stress intensity at the crack tip ΔK_{eff} as the stress intensity range that drives the crack given by:

$$\Delta K_{\text{eff}} = K_{\max} - max(K_{\min}, K_o) \tag{19}$$

where K_o is the crack opening stress intensity.

1976 EIFS Rudd and Gray

By using fractographic results and back extrapolating, Rudd and Gray [9][10][11] produced an estimate of the equivalent flaws that, when grown forward with a master curve based on the average of individual crack growth curves, would produce a distribution of cracks at some time (when the cracks are measurable) which is the same as the measured distribution of crack sizes.

5.4 Crack Growth Laws

It is traditionally recognized that some version of the Paris crack growth law (Equation (18)) coupled with a cycle counting technique [64] gives the best estimation of crack growth. This is done in essentially a cycle by cycle approach. However, when trying to account for acceleration and retardation effects, it is necessary to implement a material memory which can then modify the crack growth equation. A simple memory effect is introduced by including the R ratio in the crack growth equation. In conjunction with ΔK this means the equation has knowledge/memory of $K_{\max}$ and $K_{\min}$.

Extensions to the Paris Equation (18) were made by a number of people including Walker [65] who extended the equation to include the effect of stress ratio, and Forman [66] who modified the equation to include the effect of R ratio and accelerated growth associated with large stress intensity. The Walker equation allows for the effect of R ratio to be matched with experimental data. However, the Forman equation is somewhat more limited since the R ratio effect is coupled with the determination of the acceleration effect. Nonetheless, the Forman equation is more widely used.

The Walker equation [65] for predicting the rate of crack growth is:

$$\frac{da}{dN} = C\left[\frac{\Delta K}{(1-R)^{1-m}}\right]^n \tag{20}$$

The Forman Equation [66] is similar:

$$\frac{da}{dN} = C\frac{\Delta K^n}{(1-R)K_C - \Delta K} \tag{21}$$

The Forman equation takes into account the rapid rise of crack growth rate as the maximum stress intensity factor approaches the fracture toughness K_C. There are a number of models available to explain the stress ratio R dependency, although the physical reason for this dependency is not precisely known. Crack closure proposes some rationale, however some materials show a significant effect of negative R ratios down to -1, which cannot be explained by crack closure. There are also explanations provided by crack tip blunting, residual stresses close to the crack tip and possibly residual strain hardening.

Ultimately the determination of all crack growth rates is performed experimentally (and generally only constant amplitude data are published), however progress has been made on further identifying those features that determine the similitude conditions so that data obtained from constant amplitude tests can be applied to an arbitrary sequence of loads.

Another feature noted about crack growth is that it tends to be self similar. This means that over a large range of scales the crack looks the same. The two most important features are the size of the plastic zone ahead of the crack, and the crack length. Provided the ratio between these two features remains constant, then the crack growth rate should only be proportional to the crack length. This results in an exponential rate of crack growth which was found by Frost and Dugdale [48][67] to be in agreement with measured crack growth rates.

The constants for all these equations are generally based on constant amplitude data (although this is not a requirement). When these crack growth equations are applied to fatigue sequences experienced by aircraft they tend to result in conservative predictions by factors of the order of 2–10 [68].

In attempting to improve the prediction of crack growth, a number of researchers such as Wheeler [69] and Willenborg [70], proposed modification of the rate of crack growth based on the effect of the crack tip growing into a plastic zone left from previous overloads. Wheeler believed the retardation should be related to some fraction of the size of the plane strain plastic zone a_1, and Willenborg believed it should be related to the size of the plane stress plastic zone a_2. Measurements by Lankford and Davidson [71] showed the measured size of the plastic zone ahead of a crack after an overload, was in between the predicted plane stress and plane strain sizes for intermediate overloads, and closer to the calculated plane strain size for large overloads.

They modelled the process by the crack slowing down as it moves through the plastic zone. Actual measurements of the rate of crack growth show that there is an initial acceleration of the first few striations before they begin to decelerate and slow down for the next few hundred striations, then gradually increase in size until they resume their normal rate of growth.

In addition, they did not consider that the direction of the crack growth is influenced by the type of loading applied in the cycle. Lankford and Davidson [71] showed that the fatigue crack tends to change direction after an overload and tends to avoid the most intense zone.

The plastic zones are given by [71]:

$$\text{plane strain} : a_1 = \frac{1}{2\sqrt{2}\pi}\left(\frac{K_{ol}}{\sigma_y}\right)^2 \tag{22}$$

$$\text{plane stress}: a_2 = \frac{1}{\pi}\left(\frac{K_{ol}}{\sigma_y}\right)^2 \tag{23}$$

where σ_y is the flow stress at one percent plastic strain and K_{ol} is the overload stress intensity.

5.5 Crack Closure

In order to calculate crack growth rates, it is necessary to determine the stress intensity at the tip of the crack. This can be found for different types of cracks in particular geometries. These can best be obtained from standard handbooks [72]. However modification can be made to the standard stress intensity factor to incorporate sequence effects. In analyzing sequence effects, Marris [73] found that the rate of crack growth rate increased if compression load cycles were introduced into an otherwise zero to tension loading. A number of models have been proposed to account for retardation and acceleration. Early models were based on the plastic zone size at the tip, such as the Willenborg model [70] and the Wheeler model [69]. Present day models seem to favour crack closure [74][75].

In thin sheet material, large overloads in a loading sequence cause plastic deformation ahead of the crack. It is postulated as subsequent loads cause the crack to grow through the plastic zone caused by the overload, the stretched zone passes into the wake of the crack, leaving a lump on the crack face. These lumps created through the plastic zone process have the effect of propping open the crack and thus reducing the effect of small loads on the crack tip. This is because the crack tip is not open at any loads below a given threshold, since compression of the lump takes place. This effect is known as crack closure, and although postulated by Elber [63] as due to plasticity, it has also been observed due to surface roughness and corrosion.

Normally the crack closure is considered to be caused by plastic deformations left in the wake of the crack, but oxide debris, crack branching and crack surface micro-roughness may also cause crack closure [76].

The problem with experimentally determining the effect of plasticity induced crack closure is that experiments cannot distinguish between roughness induced closure which may result in some point contact over the surface of the crack face, and plasticity induced closure which is supposedly a uniform contact over the length of the crack tip. Another drawback of the plasticity induced crack closure model is that it treats the material as homogeneous, whereas at the scale of striation formation, the material behavior is dependent on the grain orientation, and the mechanisms at this level involve slip bands and cleavage. In addition, these mechanisms are also dependent on the environment present at the crack tip.

Allowing for the effect of the crack tip not experiencing the full range of stress intensity, goes some way to explaining the discrepancy between predicted crack growth from constant amplitude coupon data when applied to a variable spectrum. However, despite over thirty years of investigation and characterisation, it is difficult to be certain of the actual mechanism [77]. This is because it is difficult to directly study the crack tip and it is only done through indirect means. Thus there remains significant dispute as to whether crack closure exists and its effect on the rate of crack growth [78]. Some methods used to investigate crack closure are listed below:

1. striation spacing for special load sequences,
2. determining wake contact from ultrasonic diffraction, potential drop and X-ray techniques, and
3. direct measurement of the ΔK_{eff} from the stress field such as from the method of caustics.

This difficulty in defining and modelling closure is perhaps one of the reasons it is not universally implemented, even though the concept of an effective driving force on the crack tip is widely accepted.

However, fatigue crack growth models based on closure have generally shown an improvement in being able to predict crack growth of a realistic spectrum [74]. The de Koning model [74] is called CORPUS (Computation Of Retarded Propagation under Spectrum Loading). This model is based on an approximate description of opening and closure. The model calculates the crack tip opening stress for a cycle as a function of the maximum stress intensity $K_{\max}$ and the minimum stress intensity $K_{\min}$ that may come any time after $K_{\max}$, but will have the effect of reducing the size of the opening stress (large compressive loads reduce the height of the lump that props the crack tip open).

The procedure defined by the model keeps track of all the lumps on the crack face while the crack tip is within the monotonic plastic zone generated from $K_{\max}$. The largest lump will prop the crack open, but all lumps can be subsequently reduced due to subsequent underloads. There are also minor refinements in determining the size of the plastic zone and the effect of material thickness, which determines whether there is a plane stress or plane strain condition which also significantly affects the size of the plastic zone.

Two plastic zones are created, a monotonic plastic zone which is created in the material ahead of the crack tip into either pre-yielded or untouched material, and a reversed plastic zone which is much smaller (typically a quarter of the size of the monotonic zone) which is created from the compressive loads. The reversed plastic zone, when penetrated by the crack tip, creates lumps on the crack face which prop open the crack. There is little effect due to the size of the reversed plastic zone and it is only significantly effective while the crack is within the monotonic plastic zone.

By processing the sequence on a cycle by cycle basis the crack tip opening stress can be determined and so the crack growth is calculated. As per most crack growth models, the crack size is then the summation of all crack growth increments.

Wanhill [43] in a review of the ability to predict fatigue crack growth, has noted that this sensitivity to closure is more apparent for thin sheet materials where the plastic zone is larger, due to plane stress conditions, than in thicker materials. It can also lead to more variation in crack growth when the large loads are infrequent or the crack growth rates due to either material selection or loading are small, since more crack growth occurs in the plastic zone.

All these characteristics of thin sheet, infrequent large loads and low rate of crack growth are typical of transport aircraft. Thus, particular attention should be paid to these factors when dealing with a probabilistic analysis of these types of aircraft.

5.5.1 Strip Yield Models

A number of models have been developed that predict the amount of plastic deformation just ahead of and behind the crack [79][80]. Strip yield models were first developed by Hill in 1976 using Linear Elastic Fracture Mechanics (LEFM) to determine the deformation of material around a fictitious crack which includes the plastic zone. By assuming the material is elastic-perfectly plastic, the amount of plastic stretch can be calculated at different positions along the crack face. By keeping track of the stress levels required to bring the faces into contact an estimate of the crack opening stress can be made. By coupling this model with crack growth models, the cycle by cycle crack growth can be determined for an arbitrary sequence. The drawback with this method is the large computation time required to solve for the crack opening stress of each cycle.

The degree of restraint at the crack tip (whether plane stress or plane strain or some intermediate value) has been found to be an important parameter in predicting the correct level of plasticity.

5.6 Short Crack Regime

While LEFM has been successful in predicting the rate of crack growth under a variety of different loads, there have been discrepancies in what has been termed the short crack domain (typically less than 125 μm or 0.005 in). It has been observed that for short cracks the rate of crack growth is not always a simple function of the stress intensity, with significant variation being observed, and is generally higher than for long cracks.

Leis and Forte [81] attribute the short crack effect as being due to the restraint on the tip of the crack when in a plastic zone caused by a stress concentration. They indicate fracture mechanics still apply in this region but should be equivalent to the crack growth experienced by a coupon in an equivalent constant (yield) strain restraint instead of constant load.

By comparing the crack growth rate for coupons with load control with those using edge strain control they were able to show that the short crack effect is not characterised by the length of the crack but the control condition at the tip of the crack. A notch with a large plastic zone may develop a long crack which exhibits the short crack effect.

Part of the problem of dealing with short cracks is in generating unambiguous data. This is made difficult by the following sources of error in measuring a short crack:

1. locating the crack tip; surface cracks may show significant branching, making it difficult to measure the crack tip,
2. determining the absolute length of the crack,
3. determining the relative advance of the crack, and

4. there may be many small cracks before a single dominant crack develops.

5.7 What is a cycle ?

One drawback with nearly all crack growth programs or models is the lack of a precise definition of what constitutes a cycle. For constant amplitude loading it is fairly obvious to determine, however for a variable amplitude sequence it is really a question of how to estimate variable amplitude crack growth rates from constant amplitude data. There is no specific requirement that constant amplitude data must be used, but that is the data that is generally available [82].

McMillian and Pelloux [61] have shown that crack growth occurs only during the rising portion of load cycles.

5.8 Predicting crack growth from a representative sequence

Generally, direct use of the Paris equation does not produce good agreement with a variable load amplitude test result.

Barsom [83] showed (for steel) that a modification of the Paris equation (Equation (18)), also proposed by Paris, was more effective in predicting random sequences. He used

$$\frac{da}{dn} = A\Delta K_{\mathrm{RMS}}^{m} \tag{24}$$

where A and m are constants and $\Delta K_{\mathrm{RMS}} = \sqrt{\frac{\Sigma_{i=1}^{n}\Delta K_i^2}{n}}$ is the root mean square value of the stress intensity ranges in the sequence. A similar approach known as the Block Approach developed by Barter et al. [68].

5.9 Using fractographic data to determine initial flaw size

When cracks are small, most of the variability comes from the initial flaw size. Fractography can be used to measure the rate of crack growth for each repeated block if there is an easily recognizable pattern of loads within a block. This is most applicable for laboratory test pieces where a controlled sequence can be applied repeatedly. However cracks from service aircraft become more difficult to analyse because of the lack of a repeating pattern. Sometimes however, environmental effects also leave their mark on the surface of a crack (through a discoloration of the surface), such as when an aircraft is brought in for servicing. Knowledge of these events in combination with examination of the fracture surface can reveal the true rate of crack propagation.

Fractographic analysis is generally performed with an optical microscope using magnifications of 200–500 times. An optical microscope is more able to show color variation on the surface, and provides a larger view of the crack so that a number of blocks of the repeating pattern can be seen in the field of view at one time. This allows the crack front to be tracked. Too high a magnification can easily lead to losing track of the crack progression.

Because these patterns or any fatigue striations may not be easily identified when the crack growth rate is small, extrapolation of the crack growth back to zero time is required. A number of ways have been proposed to back extrapolate the fractographic curve but these are generally variations of two methods.

extrapolate back the a vs time curve This involves simply plotting the crack size (usually on a logarithmic scale) as a function of time and then extrapolating this curve back to zero time to find the equivalent initial flaw size.

extrapolate back the rate of growth da/dt versus a curve Since for a constant load the only thing changing is the crack size, the change in the rate of growth should be a function of the crack size. This can be expressed by fitting a simple equation of the form [84]

$$\frac{da}{dt} = Qa^b \tag{25}$$

in which Q and b are parameters depending on loading spectra, structural material properties etc. However any smooth fitting curve to the data should be sufficient. Because this function depends only on the difference in crack size the curve is independent of the initial flaw size. So the initial flaw has to be found by integrating this curve from a known starting point.

6 Probability Distributions Used

A probability provides a quantitative description of the likely occurrence of a particular event. Probability is conventionally expressed on a scale from 0 to 1; a rare event has a probability close to 0, a very common event has a probability close to 1.

How do I choose a distribution? There are a number of different distributions that can be used in risk calculations. Some of the more common distributions that are used are the normal, log normal, Weibull and Gumbel. In some cases the distributions fit the data perfectly. In the cases where a good fit cannot be found it becomes important to find out which section of data is most critical in the risk calculation.

For example in the case of stress exceedance data the most critical part of the data are the exceedances at the higher stresses. Therefore it is important to fit this section of the data more accurately than the lower stress exceedances.

If you have some information on the mechanism of the formation of the variables it may be possible to choose a distribution using this information. For example, a random variable for the length of a chain which is the sum of many (unknown) random length links will generally be a normal distribution (based on the central limit theorem).

The product of random steps results in a lognormal distribution. For example random steps such as the growth of an inclusion in a time step where the growth is a random factor of the previous step's inclusion size should result in a lognormal distribution.

Similarly the Weibull will best represent the distribution of the life (or strength) of the chain since the life of the chain is based on the shortest life of any link (based on extreme value theory).

Gumbel will be obtained as the maximum life or strength of the chain. Hence exceedances of random data are typically Gumbel.

The maximum likelihood method is a method to determine the parameters of best fit for a given distribution to a set of data. In the case of a normal distribution, these parameters would be the mean and standard deviation. Although the method is not based on any firm mathematical theory [85], it does produce good estimates of the unknown parameters. The method is based on the premise that the data collected will be the most probable set of all data obtained from the distribution. To work out the probability of this data set occurring is simply a matter of working out the frequency of each data point. The frequency is obtained from the PDF. So the probability of a data point occurring is proportional to the value of the PDF at that point. So the probability L (the likelihood) of the entire data set occurring is the product of the probabilities of each of the individual observations x_i multiplied together:

$$L = \prod_{i=0}^{n} f(x_i) \tag{26}$$

The maximum value of the likelihood function can be determined by performing a non-linear maximisation to vary the unknown parameters that characterise the distribution, which are say, the mean μ and the standard deviation σ. The likelihood function can also include other effects such as censored data (i.e. the test was stopped before failure of the sample) and a finite accuracy in readout data [86].

6.1 Time to failure/crack initiation

Many authors [15][87][88][89] propose the use of the Weibull distribution (two parameter) for the time to failure or the time to crack initiation. Although Yang and Trapp [90] have indicated that use of a Weibull may be unconservative.

The Weibull distribution is given by the equation:

$$f(t) = \frac{\alpha t^{\alpha-1}}{\beta^{\alpha}} \exp\left[-\left(\frac{t}{\beta}\right)^{\alpha}\right] \tag{27}$$

$$F(t) = 1 - \exp\{-(t/\beta)^{\alpha}\} \tag{28}$$

where $F(t)$ is the distribution function of the probability that the fatigue life is less than or equal to t, α is the shape parameter and β is the scale parameter. A three parameter Weibull distribution can be obtained by replacing the variable t with an offset variable $t - \gamma$. This means that no failures will occur before time γ. This can be loosely applied to fatigue failures that take some time to develop.

However Yang and Trapp [90] examined aircraft structures in service and found that the distribution of time to crack initiation is not Weibull, and that the prediction based on the Weibull distribution is unconservative in early service life.

6.2 Initial flaw size distribution

The normal distribution is given by:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^{2}\right] \tag{29}$$

$$F(t) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{t} \exp\left[\frac{1}{2}\left(\frac{\xi-\mu}{\sigma}\right)^{2}\right] d\xi \tag{30}$$

The lognormal distribution is:

$$f(t) = \frac{1}{\sigma'\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln t-\mu'}{\sigma'}\right)^{2}\right] \tag{31}$$

$$F(t) = \frac{1}{\sigma'\sqrt{2\pi}} \int_{0}^{t} \exp\left[-\frac{1}{2}\left(\frac{\ln \xi-\mu'}{\sigma'}\right)^{2}\right] d\xi \tag{32}$$

where μ' and σ' are the mean and standard deviation of $\ln t$ (not t).

Yang and Manning [91][92] proposed the Weibull compatible distribution for representing a cumulative EIFS distribution. The Weibull compatible distribution produces a Weibull distribution of time to failure when used with an exponential rate of crack growth. The distribution is given by the equation:

$$F_{a(0)}(x) = \exp\left\{-\left[\frac{\log(x_u/x)}{\phi}\right]^{\alpha}\right\} \qquad 0 < x < x_u \tag{33}$$

where $a(0)$ = EIFS and is the crack size at time $t = 0$, x_u = EIFS upper bound limit and α and ϕ are empirical parameters. Artley [13] suggests the upper limit to flaw size may be based on a lower limit of inspectability.

Gruenberg et al. [93] conducted a series of fatigue tests on notched specimens (half round hole on one side of the coupon) to compare the distribution of fatigue lives with those predicted from assuming cracks grow from inclusions. Inclusions were measured to obtain a log normal distribution of particle areas and these were then used to derive an effective radius of crack initiation. Good results were obtained for crack breakthrough to surface (i.e. subsurface initiation) and for total fatigue lives. Better results were obtained for shorter lives.

6.3 Crack growth distributions

For measuring crack growth rates it has been proposed by Artley [94] to use

$$\log(da/dt) = \log(C) + n \log \hat{K} \tag{34}$$

where $\hat{K}$ is the spectrum peak stress intensity factor. Thus this is a linear relationship on a log-log scale. Note the similarity of Equation (34) to Equation (16).

Yang and Manning [95] indicate that lognormal random variables are appropriate to use for modelling crack growth variability, simply because lognormal results in conservative crack growth predictions, as its statistical dispersion is the largest among the class of general random processes. They also suggest that it is useful because it is mathematically simple, it is conservative, and yet it accounts for the variations in material crack growth resistance, usage severity, and utilisation. Further only a small number of tests are required to calibrate probabilistic crack growth models. Reasonable crack growth predictions for structural details have previously been obtained for both coupon specimens and full scale aircraft structures.

The variability in da/dt can be accounted for using a random variable X which has a lognormal distribution by:

$$da/dt = Xg(\Delta K, R, a) \tag{35}$$

where X accounts for the overall statistical variability of the crack growth rate and because there are other sources of variability such as material crack growth resistance, usage severity, aircraft utilisation and environmental effects, these can all be accounted for with separate lognormal random variables by combining them into a single random variable $X = H_1 H_2 H_3, \ldots$, where H_i are all statistically independent lognormal random variables. Using this approach, the standard deviation of the overall crack life variability σ_X is given by:

$$\sigma_X = \sqrt{\sigma_{H_1}^2 + \sigma_{H_2}^2 + \sigma_{H_3}^2 + \ldots} \tag{36}$$

where σ_{H_i} are the standard deviations of the independent lognormal random variables. They have also found that the fatigue life consumed per flight hour from tracking of individual aircraft, follows the lognormal distribution reasonably well.

An example is given for a fighter aircraft where material variability has a $\sigma = 0.131$ and the variability due to usage severity is $\sigma = 0.249$ based on limited fleet tracking results.

In this example they define a minimum acceptable level of risk of 10^{-8} per flight hour. This is based on historical data for what appears to be a single component.

6.4 Size Effect

Early in the development of fracture mechanics and its use in predicting fatigue crack growth there was significant research on the size effect. Because the probability of failure significantly depends on the the initial defect it was thought that the greater the area or volume within the fatigue critical area the more chance there was of having a larger flaw thus leading to a shorter fatigue life.

Weibull [96] investigated the effect of different size notches on the fatigue life in coupons of 24S-T4 (now 2024-T4). However, he divided the fatigue growth into regimes as initiation which ends in a visible crack and propagation ending with a final rupture. Unfortunately, much of the crack growth occurs from micro-flaws in his initiation period. However, he did note that the duration of the initiation period was affected not only by the stress concentration but even more by the dimensions of the notch. The larger the coupon size the lower the fatigue life, given everything else is constant.

If the probability of having a flaw is constant for a given area or volume, then according to the theory of extreme values, increasing the size of the area sampled will shift the probability towards the larger size flaw. Kugel [97] found that there is a linear relation between the log of the maximum stress and the logarithm of the appropriate volume. This relationship was valid for both smooth and notched members.

There is still extensive ongoing work relating to the effect of size, though this seems to be of particular concern to the civil construction industry, where small samples of construction materials such as concrete are used to predict the overall strength of large structures such as dams. The effect of size is expected to be more significant on the outlier 'safe life' statistics than it is on the average lives.

6.5 Residual Strength

Fracture toughness is often assumed to be normally distributed, or Weibully distributed [98], whereas the two physical quantities used to determine K_{IC} are the crack length and stress which are assumed to follow exponential or lognormal and Weibull respectively. The coefficient of variation of the fracture toughness was evaluated for a number of different materials: aluminium, titanium and steel alloys such as the D6AC steel [99], which has a larger than normal scatter when compared to other steels. The coefficients of variation for a 50% confidence level were [98]: aluminium alloys 0.03, titanium 0.05, steels 0.05, D6AC 0.1. This should be applicable to all types such as plate, sheets and extrusions, with the exception of forgings. It would be expected that the scatter would be larger in forgings.

6.6 Inspections

For a risk analysis, there are two ways in which inspection intervals can be chosen. These are:

1. Use intervals provided - In most cases the analysis of an existing scenario is being performed, and thus already has an existing inspection routine. In many cases the analysis is performed to see how extending the interval will affect the risk to the aircraft.

2. Create inspection intervals - In the case where inspection intervals are not provided, the best method to choose inspection intervals is by calculating the risk. The main aim is to maintain the aircraft with the risk below the allowable risk limit, therefore an inspection should be scheduled for the time when the aircraft is at the allowable risk limit. In this manner an inspection schedule can be created.

It is generally assumed in risk assessments that the probability of detection during an inspection of a crack is purely a function of the crack size. Typically this is the crack length, but some inspections such as eddy current or ultrasound are also dependent on the crack depth. This is because the eddy currents may not be sufficiently disturbed if the crack is too shallow.

By assuming independence between two inspections A and B, the probability of detecting a crack can be calculated as:

$$PoD(a) = 1 - (1 - PoD_A(a))(1 - PoD_B(a)) \tag{37}$$

Thus this implies that repeated inspections (either performed immediately or at some later date) will always have a beneficial effect.

However Erland [100] conducted a study into the effect of repeated inspections by different inspectors on a series of different sized cracks using Fluorescent Penetrant Inspections (FPI). Erland showed that there is some dependence between inspections. This implies that merely inspecting twice will not have the full benefit as indicated by Equation (37).

Inspection data are dichotomous. Either a defect is found, or it is not. This type of process can be described by a binomial probability density function. The POD of flaw size a is then found by using a maximum likelihood approach. Maximum likelihood estimation finds the parameters of the binomial distribution that make the data observed the most probable or likely to occur. Thus the probability of each data point occurring can be obtained from the PDF. The probability of all of them occurring in a set of observations is simply the product of the probability of each sample because they are independent. The parameters for the distribution are then varied in a numerical optimisation to determine those parameters that maximise this value. This is given in more detail in Section 9.3.

Erland found that multiple inspections do not increase the inspection capability because the inspections may be performed by the same inspector or the cracks have the same location or size which influences their detectability.

The number of times a single part needs to be inspected is directly related to the variance in the system. That is the more the variability between inspectors the more the system is sensitive to these differences, the larger the number of inspections required.

The true probability obtained from subsequent inspections A and B is:

$$PoD(A \text{ or } B) = PoD(A) + PoD(B) - PoD(A \text{ and } B) \tag{38}$$

The effectiveness of repeating inspections using FPI, results in an increase of POD. Erland found that the benefits of repeated inspections is reduced for large cracks. This is most likely because if they have been missed for some reason during the first inspection the same conditions will prevail (except for ability of the inspector) that will lead to the crack being missed also for subsequent inspections.

Although Erland's study was carried out for FPI, these limitations should be applicable to all other inspections for the same reasons.

A large body of cracks were inspected. This involved 22000 inspections performed on 174 cracks by 170 different inspectors. This collection is known as the "Have Cracks" data [101]. These data were statistically analysed by Berens and Hovey [102]. They found the best fit to the probability of detection data was with the log-odds log-scale model. This is given by:

$$P(a) = \frac{\alpha a^{\beta}}{1 + \alpha a^{\beta}} \tag{39}$$

where a is the crack length and α and β are regression parameters. It was found from the analysis of these data that cracks of the same length do not have the same detection probability. A 90% probability of detection and a 95% confidence limit could not be obtained for small crack sizes as prescribed by the damage tolerance requirements of MIL-A-83444.

Berens and Hovey went on to carry out a risk assessment of an example transport aircraft using the "Have Cracks" data. In this analysis, they used a Gumbel distribution to represent the extreme stresses and they determined the probability of a crack exceeding the critical crack size within a time interval of 5000 flights. This model is based on that of Yang and Chen [103].

Although PROF includes probability of detection it does not consider any of the dependence factors discussed here, and thus assumes inspections are totally independent. This invariability leads to the conclusion that more inspections are better. The POD is modeled in PROF by the equation:

$$\mathrm{POD}(a) = \left\{ 1 + \exp - \left[\frac{\pi}{\sqrt{3}} \frac{\ln(a - a_{\mathrm{min}}) - \mu}{\sigma} \right] \right\}^{-1} \tag{40}$$

where a is the size of the crack being inspected and $a > a_{\mathrm{min}}$, μ is the natural logarithm of the median detectable crack size, i.e. the crack size detected 50% of the time, and σ is the scale parameter, with a larger μ implying a flatter POD(a) function and lower detectability at bigger crack sizes.

7 Acceptable levels of failure

Orisamolu [1] says the selection of target reliability levels is a sensitive management issue, that must be carefully carried out to ensure that human life is not endangered, while at the same time avoiding setting an unreasonably high reliability target. This would have severe economic and operational considerations. Orisamolu recommended that the Canadian Forces use the reliability levels used by the USAF, or else base them on the current reliability levels in the Canadian Forces fleet.

Payne [104] looked at the historical failure rates of early aircraft and concluded that the proposed safety condition for reliability from fatigue should have a maximum risk rate $< 10^{-7}$ per hour and an average risk rate $< 10^{-8}$ per hour for civil aircraft. Payne et al. [105] later proposed a maximum risk rate of $< 10^{-6}$ per hour for military aircraft.

The Joint Services Guide 2006 (Ref. [106]), the current airworthiness standard followed by the USAF and USN, contains rationale, lessons learnt and instructions necessary to tailor the requirements and verification task. Therein, the stated maximum acceptable frequency of structural failure leading to the loss of aircraft is 1×10^{-7} occurrences per flight.

In general, standards provide acceptable levels of failure. Acceptable levels are also quoted in the literature, although there can be some confusion concerning whether the probability of failure relates to the total probability or the instantaneous probability, and whether it applies to a single item such as a rivet, or a complete component such as a wing, or to the whole aircraft.

The usual acceptable level of failure is 10^{-7} for the single flight probability of failure of the whole aircraft. This was proposed by Lincoln [107] but follows on from a number of researchers who have proposed the threshold at similar levels. This is also similar to the acceptable system level of failure proposed by the Mil Standard of 10^{-6} per flight for critical systems. The 10^{-3} probability of failure for the cumulative probability of failure of an aircraft as specified in DEFSTAN-970 [4] produces similar instantaneous risk levels [14] as the 10^{-7} proposed by Lincoln.

Although the risk of failure is given on a per aircraft basis, in reality it perhaps should be the risk of a fatality such that the more people that are carried in an aircraft the lower should be the acceptable risk, or otherwise take into account the actual consequences of losing the aircraft which may include the loss or disruption of the aircraft fleet which may also have significant penalties.

In conversations by the author with Dr. Lincoln, he was asked how it was best to deal with situations where there were many critical locations on the aircraft such as the F/A-18. How was it possible to determine what should be an acceptable safety level for each location when only the total probability of failure was specified in the standard. He suggested that the aircraft should be zoned. Thus, if there are approximately 10 critical zones, where a single wing might be classed as a zone, then that should have a probability of failure of 10^{-8}. In general, when a risk assessment is carried out, it is found that there are probably only a few really critical locations that contribute the bulk of the risk of failure. Thus it seems a reasonable guide to set an acceptable probability of failure for any critical location as around 10^{-8}.

However, when different standards set different types of probability of failure, inevitably one will be more conservative than the other. With fighter type aircraft, the risk of failure tends to rise quite rapidly near the end of the service life of the aircraft. Thus it would seem that most of the total probability of failure is produced within the last few hundred flights. In these instances, perhaps a single flight probability of failure is the better safety metric.

DEFSTAN 970 [4] (Issue 2, Dec 1999, Part 1, Section 3, Leaflet 35) gives the following paragraph which defines the acceptable total probability of failure:

```
2.3 The factors of Table 2 have been derived using the scatter observed
when large numbers of specimens representing structural features have
been tested under realistic (spectrum) loading. When used with the
customary statistical technique, they correspond to a probability of
failure of 1 in 2000; this makes some provision for the symmetry in
geometry and loading of [aeroplane] airframes so that a probability of
1 in 1000 is achieved for individual pairs of features; the same
factors are acceptable for helicopter airframes and dynamic
components.
```

The new Joint Services Specification Guide JSSG2006 provides guidance for typical values to be used, however the actual values used are defined for each aircraft, and hence from [5]:

```
A.3.1.2  Probability of detrimental deformation and structural failure.

Only where deterministic values have no precedence or basis, a
combined load-strength probability analysis shall be conducted to
predict the risk of detrimental structural deformation and structural
failure, subject to the approval of the procuring activity.  For the
design requirements stated in this specification, the airframe shall
not experience detrimental structural deformations with a probability
of occurrence equal to or greater than 1/10^3 per flight.
Also, for these design requirements, the airframe shall not experience
the loss of adequate structural rigidity or proper structural
functioning such that flight safety is affected or suffer structural
failure leading to the loss of the air vehicle with a probability of
occurrence equal to or greater than 1/10^7 per flight.


REQUIREMENT RATIONALE  (A.3.1.2)

This requirement establishes the maximum acceptable frequency of
occurrence of detrimental deformation and structural failures that are
```

used in conjunction with combined load-strength probability analysis.

REQUIREMENT GUIDANCE (A.3.1.2)

In some instances, historically based deterministic criteria are not applicable to the specific combination of design approaches, materials, fabrication methods, usage, and maintenance for the structural element being designed. In these instances, it may not be possible to rationally arrive at an alternative deterministic criteria and a combined load-strength probability analysis is conducted to establish that the risks of detrimental structural deformation and structural failure are acceptable. The selection of the maximum acceptable frequency of occurrence of detrimental structural deformation, loss of structural functioning or structural failure can be made by examining relevant historical repair and failure rates. A maximum acceptable frequency of permanent structural deformations would be 1 x 10-5 occurrences per flight. A maximum acceptable frequency of the loss of adequate structural rigidity or proper structural functioning, or structural failure leading to the loss of the air vehicle would be 1 x 10-7 occurrences per flight.

In most cases, a combined load-strength probability analyses is only selectively used in the analysis of the structural elements for which historically based deterministic criteria are not appropriate. In these cases, a probability analysis of a highly loaded representative structural element is performed. This analysis would address all of the significant variations in load, material properties, dimensions, etc. Once the design of the element has been completed by these probabilistic means, it is usually possible to develop a set of modified deterministic criteria which, when combined with the appropriate limit and ultimate loads, would yield the same final element design. This updated criteria can then be used to design similar structural elements. In addition to establishing new design criteria, the conduct of the probability analysis also aids in gaining an increased understanding of the more important design drivers and enables an improved design to be produced.

8 Multisite damage

Wanhill [43] concluded that multisite cracking is not a problem if the structure is sufficiently durable to reach the end of its usable life with no more than 1–2 mm of crack growth at both sides of every fastener hole.

While the probabilistic and fatigue analysis tools used to analyse single site and multi site damage are essentially the same, the analysis of multisite damage is complex, due to the number of locations that may be cracked and how to deal with the interaction of these cracks, which could lead to a situation known as Multisite Damage (MSD).

A number of researchers have looked at the problem of MSD from a probabilistic approach.

Wang [44] explains the difference between MSD and a single damage:

1. cracks do not exist at all possible locations,
2. variation due to production inconsistency may be large for the increased number of details, and
3. in MSD, loss of the load carrying capability may not lead to redistribution of the load path because of the possibility of fatigue cracks in adjacent members.

Wang [108] essentially uses the methods previously discussed (i.e. fatigue predictions using crack closure, variability in crack growth rates) in his proposal for analysing MSD. But he also considers variability in geometry, loads, and initial flaw sizes (modeled as lognormal). He also considers that, with a number of holes, not all holes will be cracked. This is modeled by assuming there is a threshold below which cracking will not occur, which is given by a probability distribution. When the initial flaw size is below the threshold, cracking will not occur. The resulting probability of crack initiation is then assumed to be normally distributed.

Order statistics may be used to give information about the largest crack in this situation [108].

9 Case Studies

9.1 USAF General Approach

Ref. [109] outlines the recent USAF approach to standard probabilistic risk assessment as applied to aircraft in their fleet and how to effectively communicate the results to decision makers. The ASIP program, initially established in 1958, recommended a safe life approach to achieve safety levels set using probabilistic levels set for aircraft. However, subsequent structural failures led to use of the damage tolerance approach, which is recommended in MIL-STD-1530A, Aircraft Structural Integrity Program Requirements.

Risk analysis can provide quantitative information to answer the following types of force management related questions:

1. Are aircraft operating at unacceptable risk levels ?
2. How do inspection interval changes impact safety and cost ?
3. What is the cost of inspection and repair versus modification ?
4. When should the weapon system be retired ?

The risk analysis method is probabilistic, since probability distributions of crack sizes, maximum stress per flight and fracture toughness and non-destructive inspection technique are required.

The risk analysis approach determines the probability that a flight will produce a stress intensity factor that exceeds the material fracture toughness based on fracture mechanics principles.

The USAF has repeatedly used risk analysis to evaluate airframe structural safety and economic issues such as static strength shortfalls, cracking of in service aircraft, and the safety and economic impact associated with service life extension, inspection and the evaluation of repair versus modification.

Dr. Lincoln [110] has written extensively on the risk assessments performed by the USAF, and has mentioned examples on F-5, T-37, T-38, C-5A, C-141, and F-16 aircraft.

Ref. [23] describes the latest version of the PROF program used for risk assessment by the USAF. The program grows a distribution of cracks forward in time. PROF's original formulation was for the growth of a single crack. More complex scenarios were analysed by additional runs of the program and external post processing. The program has since been updated and it is claimed, to facilitate the analysis of widespread fatigue damage. PROF uses the now standard approach of EIFS with a deterministic crack growth curve. Also included is the effect of inspections through a POD as a function of crack size.

PROF has been used in the analysis of the B707 aircraft [18] for the JSTARS variant [111]. These were ex-transport aircraft being refitted for a new role. In order to determine the level of structural modifications necessary to deal with the fatigue damage already accrued by the airframe, an analysis using PROF was carried out. The critical area analysed was stringer 7 in the lower wing skin. Crack data were collected by a teardown of two retired aircraft. Each of the stringer holes in the critical area was examined for cracks. This provided the probability distribution of the initial flaw sizes. Note that the flaws are cracks at the time of the teardown, which was at the end of the life of each of the two sample aircraft. Although the cracks were subsequently grown back to an earlier time, the distribution of crack sizes should be reasonably accurate because of the limited degree of extrapolation. A fuller description of the analysis carried out on the JSTARS is given by Lincoln [110].

The crack length data were collected from two aircraft. One aircraft, 707-321B S/N 19266, had a final flight time of 22 533 flights and 57 382 flight hours. The other aircraft, 707-123 S/N 17635 had flown 36 359 flights and 78 416 flight hours. A lognormal distribution was used for the upper tail of the teardown crack length data. The cracks were grown back to 40 000 hours by back extrapolating the percentiles of cracks sizes using the deterministic crack growth curve. Thus the analysis for the JSTARS aircraft starts at 40 000 hours.

PROF determines the single flight probability of failure (SFPOF). However this term is somewhat inaccurate. The SFPOF is determined by starting out with an initial flaw size distribution which is grown forward to a point of time T. The SFPOF is given as the probability of an aircraft, which then flies with this distribution of cracks to time T, failing within the next flight. This is not strictly the SFPOF, since as the aircraft ages, some of these cracks could have led to failure earlier on, but in the analysis these are all kept intact. This means the probability of failure will be gradually approaching 1 for a very old aircraft. Now the instantaneous probability of failure or hazard, should not approach 1. The PROF SFPOF then approximates the total probability of failure at long lives.

These problems with PROF were discussed more fully in the demonstrator risk analysis of the F/A-18 [14][27].

The program has later been modified to allow for the incorporation of MSD. In the original analysis [111] it was found in the original analysis of the JSTARS aircraft, that it did not meet the safe life requirement of SFPOF of 10^{-7} at 57 000 hours. However, the analysis did not seem to correctly take into account the number of fasteners involved at the critical location. The analysis appears to be for a single fastener in the aircraft. However, the more fasteners, the higher the risk of failure becomes. Typically there are about 5 000 fasteners in the critical section of the wing of the B707 (which makes 10 000 fasteners per aircraft).

Some lessons learnt from the literature will now be presented.

9.2 KC135

A number of risk assessments have been performed on the Boeing 707 aircraft [19][112] [18][113]. The initial assessment performed for the USAF [111] was to determine whether critical sections of the lower wings would result in failure when ex-airline aircraft were modified to carry out radar operations. This was the JSTARS program. To provide a distribution of flaw size data for the risk assessment, a teardown of the critical lower wing section of two retired aircraft was made. This provided a distribution of flaw sizes for each fastener in the critical section. The risk analysis was performed using PROF and followed the standard approach. The usual assumptions with the PROF analysis were made in that the distribution was for a crack at a hole and did not consider the failure of the approximately 10 000 holes in the critical section of the two wing halves. Thus the risk assessment was for a hole, not a wing.

Refs. [114][89], describing the life assessment of the KC135 aircraft, which is a corrosion prone aircraft based on the B707, came up with three metrics that are useful for judging the retirement time of an aircraft. They were:

1. safety,
2. cost per flying hour, and
3. availability.

The risk analyses currently carried out generally only address the safety issue, but clearly other factors can determine the life of the aircraft.

9.3 F16

The risk assessment approach the Israeli Air Force used for the F16 was originally conducted using a Weibull analysis [88]. The Israeli Air Force chose this approach because it:

1. has broad applicability,
2. has a simple graphical solution,
3. can be useful when data are insufficient such as when there is a small sample size, and
4. usually provides the best fit to the type of data encountered in structural failure [87].

Using this approach the analysis used a three parameter Weibull probability distribution to model time to failure data collected from field failures. The third parameter was a time offset that marked the start of possible failure and possibly corresponded to a crack initiation time before crack growth occurs.

Estimated probabilities of failure for a failure position of rank i out of a sample size N, can be made from the median cumulative function:

$$F(t) = \frac{i - 0.3}{N + 0.4} \tag{41}$$

for comparison with the determined Weibull distribution. The parameters for the Weibull distribution are determined by calculating the Weibull slope parameter β_{cr} and the characterised failure time η_{cr}. The analysis used the maximum likelihood method for the estimation of the two parameters. It was initially assumed the time offset was zero:

$$F_{cr} = 1 - \exp\left\{\left[\frac{(t - t_{0_{cr}})}{(\eta_{cr} - t_{0_{cr}})}\right]^{\beta}_{cr}\right\} \tag{42}$$

In the analysis it was noted that failure modes which are different cannot be plotted on the same Weibull plot. They found that for small failure samples the maximum likelihood approach tends to over estimate the βs and under estimate the ηs. However, because these effects can cancel out they conclude that maximum likelihood estimates are ultimately more accurate than rank regression (curve fitting approach) to obtain the distribution parameters.

The third parameter for the Weibull distribution is determined by an iterative method using the initial estimates of β_{cr} and η_{cr}. The parameter $t_{0_{cr}}$ was then estimated by the formula:

$$t_{0_{cr}} = t_2 - \frac{(t_3 - t_2) * (t_2 - t_1)}{(t_3 - t_2) - (t_2 - t_1)} \tag{43}$$

where t_1 is the first failure time, t_3 is the last failure time, and t_2 is the time corresponding to the linear halfway distance on the Weibull plot vertical axis.

The equation is then iteratively solved for β_{cr} and η_{cr} until there is convergence.

However, one of the downsides to using a strictly statistical approach, such as in modelling only the time to failure, is that in reality there is usually more data available that describes the failure, such as type of initiation, if it is from corrosion or material defects, and crack lengths. Thus extra information from existing failures is generally easier to come by than from additional failures. This would indicate that a more mechanistic approach in determining the failures would be worthwhile. Thus, using a Weibull analysis does not make use of all the available data from typical aircraft failures.

9.4 De Havilland Venom

Ref. [115] looks at the De Havilland Venom, of which 250 were built between 1953 and 1958. The safe life for this aircraft was originally established by three aircraft subjected to a constant amplitude loading, and from four aircraft subjected to program loads. As a result of this testing, a scatter factor of 2–3 was used to factor the test lives, giving a safe life of around 500 hours of typical Swiss usage.

Service experience provided data on the time to first detection of a crack in the lower wing. Of the 463 wings in service, fatigue cracks were found in 149 wings, with the remaining wings repaired before any detectable cracks were located.

The time to first detectable crack was plotted on both lognormal and log extremal probability paper. If the data produces a straight line, when plotted on probability paper for a certain distribution, it indicates the data is from the given distribution. In this case, the fit was slightly better for the log normal distribution, but both distributions overall seemed equally applicable. The log extremal gave a safe life (i.e. time to first failure) at 0.1% probability of failure of 144 flights and the lognormal gave a life of 188 flights. However separate authors Weibull [57], Freudenthal [16] and Gumbel [116], all suggest that there are some fundamental reasons for proposing that an extreme value distribution Type 3 (Weibull) should be used.

It is interesting that even with this large amount of actual data which comes from fleet usage, there is still a discrepancy whereby the safe life established from service data is about one third of the safe life established from test data [115]. This indicates that for low probabilities of failure, the safe life for an aircraft is highly sensitive to the input data, and it is hard to be definitive of a hard and fast safe life.

9.5 Lug and Stiffened panel

Artley [13] provides two examples of a probabilistic risk assessment as applied to aircraft components. She proposes the standard approach (Yang [117]) for calculating the probability of failure by including the effect of EIFS, stochastic crack growth and probability of detection. A lug is analysed using both deterministic and stochastic crack growth, and a stiffened panel is analysed. The numerical integrals for the probability are calculated (though not explicitly given) for the two examples.

9.6 C-141

Ref. [2] presents a review of a risk analysis performed by Lockheed Martin (LMASC) on some of the critical wing sections of the C-141 aircraft. The C-141 is a transport aircraft made from 7075-T6 in the 1960s. The aircraft were approaching the end of their life and so a risk assessment was performed to highlight areas that needed to be addressed in order to reach the full life. Two of the questions the risk analysis was supposed to answer were, "What is the risk of flying beyond 45 000 damage hours (approximately one life time)" and "Does the aircraft still have the inherent fail-safety that stands behind the slow crack growth method".

A number of areas were examined on the wing and fuselage, but no indication was given of what the acceptable risk level for each individual area was. The analysis was based on defining the risk as "the probability of losing a major component or components which could cause catastrophic failure". However, they recognized that the risk analysis was not exact because of the uncertainty in some of the input parameters such as estimated flight hours and inspections. So judgments were made about the validity of the results.

A fatigue test had been conducted on the aircraft to nearly three lifetimes, however a fleet failure occurred in an area where the fatigue cracks had not been detected during the fatigue test. In addition, the cracks occurred at the end of a large metal repair.

Lockheed Martin have developed their own program called DICE which uses the same methodology as provided in the RISKY program developed by Lincoln [118]. PROF is similar to the DICE program but uses a different method of inputting loads.

This risk analysis used the standard stochastic inputs of crack distributions, load exceedances, and residual strength.

9.7 Follow-on C-141 risk analysis for multi-element damage

The previous risk analysis covered the calculation of the risk of failure of a single component, although it may indeed be a singe rivet hole instead of the whole component. However, the critical wing splice joint is made up of three separate components. A new risk assessment performed by Berens, Gallagher and Dhar [119] is a follow-on to the risk analysis performed by Lockheed [2] on the C-141 aircraft.

This report tried to consider multi-element damage and MSD. Multi-element damage is when cracks cause the structure to lose its ability to sustain an increase in stress level due to load shedding, whereas MSD is when the cracks are close enough to affect the growth of each other. They use PROF in this analysis because they wanted to take into account a number of elements. However, they had to use PROF with multiple runs to get the appropriate data out.

The wing splice is made up of three elements. The chordwise joint, the splice fitting and the beam cap. To work out when failure occurs, they developed a fault tree which showed the paths of failure that incorporated each of the elements and how its status affected the failure of the joint as a whole. They said that the fault tree was to demonstrate that the wing joint failure probability could be modeled as a weighted average of the probability of fracture of the chordwise joint, given the status of the wing splice fitting and the beam cap. The weighting factors are the probabilities of fracture of the splice fittings and the beam cap. Thus the formula they give for probability of failure is:

$$\begin{aligned}
POF &= \Pr(CSF, SFTAC, BCTAC) + \Pr(CSF, SFTAC, BCF) \\
&\quad + \Pr(CSF, SFF, BCTAC) + \Pr(CSF, SFF, BCF) \\
&= \Pr(CSF|SFTAC, BCTAC)\Pr(SFTAC)\Pr(BCTAC) \\
&\quad + \Pr(CSF|SFTAC, BCF)\Pr(SFTAC)\Pr(BCF) \\
&\quad + \Pr(CSF|SFF, BCTAC)\Pr(SFF)\Pr(BCTAC) \\
&\quad + \Pr(CSF|SFF, BCF)\Pr(SFF)\Pr(BCF)
\end{aligned} \tag{44}$$

where

CSF	= chordwise joint fracture
SFTAC	= splice fitting intact
SFF	= splice fitting fractured
BCTAC	= beam cap intact
BCF	= beam cap fractured
$\Pr(A, B, C)$	= Probability of events A and B and C
	= $\Pr(A\|B, C)$ $\Pr(B)$ $\Pr(C)$
$\Pr(A\|B, C)$	= conditional probability of event A given the events B and C

They note that the analysis performed is only for one location on the wing, however there are two wings on each aircraft. Based on this they give the formula for the probability of failure of the aircraft:

$$\mathrm{POF}_{\mathrm{A/C}} = 1 - (1 - \mathrm{POF})^2 \tag{45}$$

There are also about 265 airframes in the fleet for which the analysis is applicable. Thus they give the probability of having a failure in the fleet as:

$$\mathrm{POF}_{\mathrm{fleet}} = 1 - (1 - \mathrm{POF}_{\mathrm{A/C}})^{265} \tag{46}$$

However the formula used for the risk of failure in the aircraft is not strictly correct, since the probability of failure is made up of the probability of having a crack and the probability of having a load large enough to cause failure. Now, the probability of having a crack is independent for all locations (since the initial flaws are very much smaller than the distance between the cracks). However, the probability of having a load big enough to cause failure will be the same, since the wing experiences a single distributed load at any given instant. This means that the total probability of failure for a particular item is not independent of the other locations on the aircraft. So the formula given is strictly incorrect. The formula they give for the fleet is correct since each aircraft will see independent loads. No mention is made of this in the report, although the error is small for this analysis.

In addition, the analysis is based not on a joint but a single rivet in the joint which may contain many hundreds. If you apply the same formula to work out the failure of the joint or the aircraft consisting of two joints, a significant error will be made. Yet the analysis in the report does not mention the number of rivets in the component they are analysing.

The original data were obtained from Lockheed. To determine the effect of inspections, they also included a distribution of cracks that represent the population of repaired cracks once they were detected by inspection. Essentially the repair quality replaces the detected cracks with a very small initiated crack. Thus, because these cracks were so small, the actual values of the distribution were found not to affect the analysis.

They also used standard values for the fracture toughness variability of 5% coefficient of variation. They found that the probability of fracture was not sensitive to the assumed coefficient of variation. This is because the rate of crack growth is sufficiently high that given enough time fracture will occur with even a fairly frequent load load.

Separate runs of PROF were obtained for each of the elements in the joint. Crack growth curves were determined for each of these elements with appropriate combinations of the different elements broken.

The analysis initially had crack growth curves with large jumps in them. This was caused by cracks crossing holes. The PROF manual [22] indicates that in principle such jumps should be able to be successfully modeled in PROF. However, the jumps had to be removed because they created numerical integration inconsistencies.

The crack size distributions were derived to model the cracks at 31 000 spectrum hours, which was the zero reference time for the risk analysis. At this time 70% of the beam cracks had a size greater than 1 inch. Thus more than 70% of the beam caps were in a failed state at the start of the analysis. The original PROF code had to be modified to handle the integration over these large crack sizes.

The initial inspection period was set at 328 hours. The risk analysis was performed both with and without an inspection at the reference time of 31 000 hours. The average flight length was 3.5 hours. They obtained the POD data from the “Have Cracks” data set [101].

PROF computes the single flight probability of fracture at approximately 10 equally spaced intervals throughout each usage interval. The usage interval is specified in terms of spectrum hours from the zero reference time, and defines the times at which inspection and repair actions occur. PROF also calculates the interval probability of fracture but only at the end of the usage interval. They found that the single flight and the interval probability of fracture were equal to within three significant figures for the location analysed.

To work out the probability of failure they required the interval probability of failure. However since this was not readily available they used the SFPOF because they believed it to be the same. The analysis indicates that the equivalence of the SFPOF and the interval probability of failure is a result of the fracture probabilities of these elements being driven by the percentage of cracks reaching unstable size rather than the occurrence of a single large stress.

Previous work with PROF by the author of this report has shown that, because of the formulation of the SFPOF in PROF, the program starts out calculating the SFPOF, but over some time, because of the approximate nature of the formula, it essentially transitions to approach the total or interval probability of failure. It seems, because of the large crack sizes being analysed [119], the time of 328 hours was long enough for the SFPOF to transition to the cumulative probability of failure. Thus it was believed in the analysis that it was the same throughout the entire interval. However this should not be the case.

Initially at 31 000 hours, the program started off with a very high fracture probability of the beam caps. This was due to the large proportion of very large cracks in the initial conditions. However, after the first inspection and repair cycle at 328 hours, a large number of these cracks were repaired and greatly reduced the fracture probability for this element. It was noted [119] that the effect of the failed beam cap on the fracture probability of the splice fitting was relatively minor in comparison to other effects.

In summary, the authors claim they were able to use PROF to predict multi-element fatigue by performing multiple runs. They felt that this process needed to be automated, so they recommended that a post processor be used to combine the results of the individual runs as part of a multi-element risk analysis. They also recommended that the PROF code should be modified to calculate the interval probabilities at intermediate times even though it would be expected to cause a large increase in computer time to perform these calculations.

9.8 B1-B Bomber risk analysis and methods

The B1-B bomber is an aging swing wing bomber with similar construction to the F-111. A critical piece of structure is the wing carry-through box. This is assembled with TaperLok fasteners to improve the fatigue life of the article. However the effect of the interference fit fasteners has not been included in the analysis used for setting inspection intervals, because no credible crack growth curve was available. Testing of components with TaperLok fasteners showed an improvement in the safe life of 9-10.

At the time there was no reliable inspection technique available that would detect cracks in the holes of the TaperLoks without removing the fastener. This meant inspection would be very costly. As the aircraft was approaching the time for inspections in 1999, a risk analysis was performed by Torng [120] to see if the inspections could be delayed.

A new technique was developed for solving the risk of failure, but it was based on the same formulation as used by PROF, because the USAF had already accepted that approach. The reason for the different solution method of the PROF formulation was given as the large coefficient of variation of the crack size distribution, an extremely nonlinear failure function, and an extremely small probability of failure. It has been noted previously that, because of the hard wired nature of the subroutines in PROF, the solution is not very accurate for single flight probabilities below 10^{-10}. Torng [120] solved for the risk of failure by using a robust importance sampling method. This used a non-gradient based most probable search scheme to locate the most probable point. (The most probable point is the mean failure point i.e. the configuration of variables that is most likely to occur in the failure population at the time of interest.) An adaptive importance sampling scheme was then used to calculate the probability of failure.

The original analysis of the wing carry through structure involved calculating the crack growth curve from an initial rogue flaw of 0.05 inches with no account being taken for the benefit of the TaperLok fasteners.

The risk analysis used a smaller equivalent initial flaw size to take into account the effect of the TaperLok fastener. They somehow derived an EIFS of 0.005 inches, though the justification for this was not provided. However, it agrees with the EIFS value used by DSTO in analysing an F-111 wing failure from a TaperLok fastener [121].

For a detailed risk assessment, they wanted enough data to provide a crack size distribution to cover the range of -6σ to $+6\sigma$ without interpolation.

Torng defined the limit state function (the function that needs to be evaluated such that failure is the part of the curve that is below 0) as:

$$g = \frac{K_c}{[\sqrt{a\pi}\beta(a)]} - \sigma_{\max} \tag{47}$$

The risk of failure P_f is given as:

$$P_f = \Pr(g < 0) = \Pr \frac{K_c}{[\sqrt{a\pi}\beta(a)]} < \sigma_{\max} \tag{48}$$

Because the goal of the analysis was to determine the risk for not performing inspections, the POD and repair distribution issues were not considered.

Torng compared the robust importance sampling method developed for this analysis with other methods. This method required only 3522 g-function calculations compared to 10^6 for sampling methods and 10^9 for a Monte Carlo method.

In calculating for the total probability of failure for the entire aircraft they calculate the probability of failure of a single location P_s. They note there are 200 similar locations in the aircraft and find the probability of failure as $200 \times P_s$, which of course is incorrect. This has been discussed a number of times previously in this report.

As a result of the risk analysis they recommended that the safety factor of the standard damage tolerance analysis be changed from 50% (i.e. divide the calculated inspection interval by 2) to 80%. Of course this little change to the standard approach is essentially not using the standard approach at all. Thus essentially the safety of the aircraft is based on the risk analysis.

A number of results for different problems were given by Torng [120]. These would make good example problems for gaining experience for analysts new to the field. A number of times the deficiencies with the PROF implementation are pointed out, such as the use of tables for the crack growth, which do not allow for sufficient refinement when the cracks are small, to allow for the accuracy required. Also pointed out is that the accuracy of the risk solution is not good when the risk of failure is small. They also point out that because the formulation of the total probability of failure and the SFPOF is similar, the results should be similar. This has also been discussed before in this report.

9.9 P-3C Reliability and Risk Analysis

Hoffman [122] performed a risk analysis on the P-3C aircraft to determine the risk of failure after the safe life had been reached. Thus the starting point for the risk analysis was the assumption of the distribution of crack sizes at the safe life of the aircraft. Since the aircraft are individually tracked and the life is calculated based on Fatigue Life Expended

(FLE) usage, the distribution of the cracks at an FLE of 1 was obtained by assuming the distribution of time to failure at the end of the safe life was lognormally distributed. It was then assumed that the loading history was known for the next 25% of life where-after the loads were assumed random. In addition to the starting crack size, the strength (critical crack size) was assumed to be randomly distributed. Cracks were then grown to determine the probability of failure.

The probability density distributions of crack sizes were updated using Baysian estimation to allow for the effect of any cracks detected from inspections.

No details of acceptable safety levels were given. The analysis appears only to be illustrative and not sufficient for fleet usage.

The general expression for the reliability of a system is given as:

$$R(t) = e^{-\int \lambda(t)\, dt} \tag{49}$$

where:

$R(t)$ is the reliability or probability of survival of the system through time t,

$\lambda(t)$ is the hazard rate or probability that a failure will occur in the next instant of time assuming previous survival.

The cumulative probability of failure in an interval dt is:

$$F(t + dt|t > t) = \frac{F(t + dt) - F(t)}{1 - F(t)} \tag{50}$$

where $F(t + dt|t > t)$ is the probability of failure before time $t+ dt$ assuming non-failure before time t, $F(t+ dt)$ is the probability of failure before time $t+ dt$, and $F(t)$ is the probability of failure before time t. The reliability function $R(t)$ gives the probability components surviving, it can therefore be expressed as $R(t) = 1 - F(t)$.

9.10 F/A-18

Boeing [123] conducted a risk assessment of the F/A-18 to predict the distribution of lives of the USN F/A-18 fleet. This was not based on individual tracking, but evaluated the probability that an aircraft would safely achieve 6000 hours of flying. It used variations calculated in the fleet usage from the Maintenance Signal and Data Recording System (MSDRS) tracking system and an estimate of the variability in the safe life of the aircraft to determine the probability of failure distribution.

The scatter in material fatigue behavior was taken from a previous evaluation of a large collection of coupon and product tests to arrive at shape factors to be used with the Weibull distribution. These were found to broadly agree with the results obtained from testing of specific F/A-18E/F components. The scatter was expressed in the variability of the time to failure. Previous work [124] looked at over 12 000 specimens and over 1000 fatigue spectrum tests to determine the probability distribution of time to failure. The Weibull probability distribution function was found to provide the best fit.

This work resulted in a shape parameter α in the range of 4.91–5.50 for aluminium alloys under spectrum loading. They found that constant amplitude testing showed more scatter that spectrum loading. This was attributed to the high loads in the spectrum failing the coupon in low cycle fatigue, whereas the constant amplitude loading resulted in high cycle fatigue. The mechanistic difference that really causes this effect was not discussed.

Using a shape parameter of 5.27, which they argued was applicable to a fighter aircraft, gave 97% of aircraft able to exceed their design life. With a shape parameter of 3.8, 90% of the aircraft exceeded the design life. This latter result was comparable to the life obtained by using a traditional scatter factor of 2 in combination with the individual aircraft tracking results. Interestingly, the 10% failure rate was not seen as unreasonable.

9.11 B707

As a result of the teardown data which found many cracks in the lower wing skin, the RAAF asked DSTO to perform a risk assessment on this area to see if any of their aircraft would be susceptible to MSD. Two risk assessments were performed on this area. The initial assessment performed by Tong et al. [18] considered the lead crack in the wing. This approach needed revising because of the requirement for a different inspection technique in the wing. The critical section of the wing consisted of four stringers, with two of them being splicing stringers that joined two separated wing skin plates together. To prevent fuel leakage from the joint a large amount of sealant was used. However this covered a large number of rivets, making detailed eddy-current inspection around each of the fasteners inside the wing a prohibitively expensive and time consuming exercise.

Since removing the sealant would be a very time consuming task, a revised risk assessment by Dixon [19] was made, allowing for the fact that the splicing stringers would only be inspected from the outside using low frequency eddy current and the intermediate stringers would be inspected internally using high frequency eddy-current around the fastener. This risk assessment was very detailed and used data from a more recent sampling inspection that the RAAF had made on their aircraft.

To analyse the wing, the Dixon risk assessment analysed two distinct sections – an inspectable section and a non-inspectable section. The inboard section, which was inboard of the fuselage and could not be accessed and hence not inspected, used an analysis of the largest crack in the section based on crack length distribution data from the outboard inspectable section. For independent cracks that do not influence each other, the probability of failure is not the combined probability of failure of each of the cracks. This is a mistake which is commonly otherwise made.

The risk of failure is essentially a product of the probability of having a crack of a certain size, and the probability of having a load large enough to fail a crack of that size. For different holes the cracks sizes are essentially independent, however the load which will cause them to break is not. Thus the probabilities can only be multiplied if the events are independent. Because of this it can be seen, that if all cracks in the region are subject to the same load and stress, only the largest crack will break. The remaining cracks contribute nothing directly to the probability of failure (unless they interact).

The distribution of crack sizes for the largest crack can be determined using:

$$F(a_{\text{largest}}) = F^n(a) \tag{51}$$

where $F(a_{\text{largest}})$ is the cumulative probability distribution of the largest crack size in a population of size n, and $F(a)$ is the cumulative distribution of the size of a crack for each location.

It can be seen from this that it is important to know the number of holes that could possibly contain a crack. So, although the numerous cracks do not contribute directly to the probability of failure, having more holes that can crack, does increase the size of the largest crack that is likely to occur. The largest crack will then be the one that fails. Interestingly, for risk analysis without inspection it is only necessary to use the distribution for the largest crack.

When inspections are included in the risk analysis it is quite possible for the largest crack to be detected. However, the remaining cracks then become important, since any one of these could grow and result in failure.

Because of the complications with inspections, the analysis was performed using a Monte-Carlo simulation. This method was chosen because the time for the Monte-Carlo approach was insensitive to the number of variables considered, whereas the solution time increases significantly when using more variables in direct numerical integration. This allowed a series of (typically 100 000) deterministic trials to be performed to estimate the rate at which cracks would lead to failure in the fleet. From this the interval probability for failure could be obtained of the increase in risk from the time of the inspection. By differentiating this with respect to time and using:

$$h(x) = \frac{f(x)}{1 - F(x)} \tag{52}$$

the hazard or instantaneous risk was able to be calculated.

Using the criteria of JSSG [106] which specifies an acceptable level of per flight risk of 10^{-7} an appropriate inspection interval could be determined.

A similar analysis was performed on the upper skin splice plate [112]. However, because of the close proximity of numerous fasteners, this analysis accounted for the interaction of the adjacent cracks on the lead crack. Crack growth was determined using the simple Forman equation (Equation (21)). However, the geometry factor was modified by compounding the geometry factor (as provided by Rooke and Cartwright [72]) of each of the individual cracks to find the effect of a number (typically two) of cracks on each side of the main crack being analysed.

The probability of failure was expressed by the equation:

$$\Pr(\text{failure}) = \int \Pr(x)\Pr(R(x) < S)dx \tag{53}$$

Thus the aircraft was assumed to fail whenever the splice plate was unable to sustain the limit load condition. The aircraft would fail when the residual strength R for configuration x was below the limit stress S. The value of the probability of failure was determined by Monte-Carlo simulation.

In this analysis, use was made of a lower critical stress than is typically found when calculating the failure stress based on allowable stress intensity alone. It was found experimentally that the stress at which cracks link up is lower. This was also demonstrated experimentally by Molent et al. [125][126]. This early failure is possibly due to the plastic zones ahead of each crack linking up. This has been accounted for by a number of different criteria, the first originally proposed by Swift [127]. A more accurate criterion was developed by Smith et al. [128] from experimental data for the 7075-T6 sheet material which was being analysed.

Because only the cracks around the largest cracks were being considered, a generalised extreme value distribution was used to represent the distribution of crack sizes for the largest crack. This is of the form:

$$G(x) = \exp[-1 + s(x-a)/b^{-1/s}] \tag{54}$$

A lognormal distribution was used for the crack sizes adjacent to the largest crack.

10 Conclusion

There seems to be ample evidence that reliability models do represent within reasonable accuracy, the fatigue process. However, one of the biggest flaws appears to be in the treatment of correlation between the random parameters. For example, in inspections, whether repeated inspections are having the same independent benefit as a single inspection, or whether other issues such as the orientation of the crack or location and accessibility of the component, reduce the effectiveness of the inspection and perhaps the added value of repeated inspections. If this correlation issue can be overcome, the reliability models may be acceptable.

The choice of distributions to use when implementing a risk assessment has surprisingly not been seen as such a critical issue outside of DSTO. There seems to be some consensus that where there is a physical underlying basis for choosing a distribution, then it is preferable to use that, rather than simply choosing the distribution of best fit. The physical conditions that give rise to a certain probability distribution have been discussed, ranging from the basis of the normal and lognormal distribution that are essentially derived when there is some sort of random cumulative process, through to the extreme value distributions which occur when the property of interest results as the largest or smallest of a collection. It should be noted that there are three types of extreme value distributions, but these occur for both maximums and minimums, and because of the asymmetry the user needs to be careful to choose the correct one.

The use of PROF also causes some difficulty. Although developed under the auspices of the USAF, it carries the status of a standard. There is strictly speaking nothing wrong or incorrect about the program or its implementation. However the formulation of the single flight probability of failure is in fact different from the accepted form. While under certain circumstances this has little effect, under other circumstances it can produce large differences between what is expected and what it delivers. In addition, the assumption of independence of critical locations is generally invalid. Again, while this makes little difference when there are only a few risk locations, when there is a large number of risk locations, as say for each rivet in the critical section of a wing, this can have a large error in the calculation of the risk. Although the calculation is conservative, this is counteracted by the fact that people in general are calculating the risk for only one critical location such as a fastener and forget to deal with the remaining fasteners.

The design standards provide acceptable levels of risk that have been essentially constant over many years.

References

1. I. R. Orisamolu. Strategic plan for the application of probabilistic methodologies to CF aircraft fleet management. Technical Report TR-95-22, Martec, Ottawa, April 1996.

2. R. P. Bell, K. M. Jones, and R. E. Alford. Thoughts on risk analysis of an aging aircraft. In *1996 USAF Structural Integrity Program Conference*, 1996.

3. S. Suresh. *Fatigue of Materials*. Cambridge Solid State Science Series. Cambridge University Press, 1991.

4. UK Ministry of Defence. *Design and Airworthiness Requirements for Service Aircraft, Volume 1 - Aeroplanes, Defence Standard 00-970*, December 1983.

5. Department of Defense. *Joint Service Specification Guide - Aircraft Structures*, October 1998.

6. *MIL-A-8860B Airplane strength and rigidity, general specifications for* , May 1987.

7. *MIL-A-87221 Aircraft structures, general specification for*, February 1985. USAF standard.

8. A. M. Freudenthal. The scatter factor in the reliability assessment of aircraft structures. *Journal of Aircraft*, January 1977.

9. J. L. Rudd and T. D. Gray. Equivalent initial quality method. Technical Report AFFDL-TM-76-83, U.S. Air Force Flight Dynamics Laboratory, 1976.

10. J. L. Rudd and T. D. Gray. Quantification of fastener hole quality. In *Proceedings, 18th AIAA/ASME/SAE Structures, Structural Dynamics and Materials Conference*, 1977.

11. J. L. Rudd and T. D. Gray. Quantification of fastener hole quality. *Journal of Aircraft*, 15(3):143–147, March 1978.

12. S. D. Manning and J. N. Yang. Advanced durability analysis - Vol I - analytical methods. Technical Report AFFDL-TR-86-3017, U.S. Air Force Flight Dynamics Laboratory, November 1986.

13. M. E. Artley. Probabilistic damage tolerance method for metallic aerospace structure. Technical Report WRDC-TR-89-3093, U.S. Air Force Flight Dynamics Laboratory, September 1989.

14. P. White, S. A. Barter, and R. N. Braemar. Preliminary structural risk assessment of RAAF F/A-18 centre fuselage. Technical Report DSTO-TR-1326, Defence Science and Technology Organisation, April 2005.

15. W. Weibull. A statistical representation of fatigue failure in solids. *Acta Polytech*, 1(9), 1949.

16. A. M. Freudenthal and E. J. Gumbel. On statistical interpretation of fatigue tests. In *Proceedings of the Royal Aeronautical Society*, volume 216, 1953.

17. A. M. Freudenthal. Reliability aspects based on time to the first failure. In *International Committee on Aeronautical Fatigue. 5th Symposium Melbourne, Victoria*, December 1967.

18. C. Tong, R. Amaratunga, A.D. Graham, and P. White. Probabilistic assessment of the failure of fastener holes in wing stringers of RAAF B707 aircraft. Technical Report DSTO-TR-1161, Defence Science and Technology Organisation, 2001.

19. B. Dixon. Structural risk assessment of RAAF B707 lower wing stringers. Technical Report DSTO-TR-1741, Defence Science and Technology Organisation, 2005.

20. A. D. Graham and G. D. Mallinson. NERF - A computer program for the numerical evaluation of reliability functions user manual. Technical Report ARL-STRUC-R-411, Defence Science and Technology Organisation, February 1984.

21. A. P. Berens, P. W. Hovey, and D. A. Skinn. Risk analysis for aging aircraft fleets Volume 1 - Analysis. Technical Report WL-TR-91-3066 OH 45433-6553, Flight Dynamic Directorate, Wright laboratory, Air Force Systems Command, Wright-Patterson Air Force Base, October 1991.

22. P. W. Hovey, A. P. Berens, and J. S. Loomis. Update of the probability of fracture (PROF) computer program for aging aircraft risk analysis. Technical Report AFRL-VA-WP-TR-1999-3030, Air Force Research Laboratory, Wright-Patterson Air Force Base, November 1998.

23. P. W. Hovey, A. P. Berens, and J. S. Loomis. Probabilistic risk assessment using an updated PROF. In *Aircraft Structural Integrity Program*, 1998.

24. C. J. Dinkeloo and M. S. Moran. Structural Area Inspection Frequency Evaluation (SAIFE) - Vol I Executive summary. Technical Report FAA-RD-78-29, U.S. Department of Transportion - Federal Aviation Administration, April 1978.

25. A. E. Anderjaska. A plan for developing an aircraft structural inspection program. Master's thesis, School of Engineering and Applied Science, George Washington University, February 1973.

26. K. D. Raithby. A comparison of predicted and achieved fatigue lives of aircraft structures. Technical Report Technical Note Structures-301, Royal Aircraft Establishment, Ministry of Aviation, London, July 1961.

27. P. White, L. Molent, and S. A. Barter. Interpreting fatigue test results using a probabilistic fracture approach. *International Journal of Fatigue*, 27:752–767, 2005.

28. A. M. Freudenthal. The expected time to first failure. Technical Report AFML-TR-66-37, Air Force Materials Laboratory, Wright Patterson Air Force Base, Ohio, February 1966.

29. C. J. Dinkeloo and M. S. Moran. Structural Area Inspection Frequency Evaluation (SAIFE) - Vol V Program demonstration. Technical Report FAA-RD-78-29, U.S. Department of Transportion - Federal Aviation Administration, April 1978.

30. M. D. Coffin and C. F. Tiffany. New Air Force requirements for structural safety, durability, and life management. *Journal of Aircraft*, February 1976.

31. H. Switzky. Designing for structural reliability. *Journal of Aircraft*, 2(6):509–516, Nov–Dec 1965.

32. P. J. Caruso. Probability of failure analysis for fracture critical F-22 titanium castings. In T. Brussat, editor, *2001 USAF Aircraft Structural Integrity*, 2001.

33. S. R. Varanasis and I. C. Whittaker. Structure reliability prediction method considering crack growth and residual strength. In *Fatigue crack growth under spectrum loads - Thirteenth Conference*, number 595 in ASTM STP, pages 292–305. American Society for Testing and Materials, 1976.

34. J. W. Provan. Probabilistic approaches to the material reliability of fracture-sensitive structures. In J. W. Provan, editor, *Probabilistic fracture mechanics and reliability*, pages 1–45. Dordercht, The Netherlands: Martinus Nijhoff, 1987.

35. J. W. Lincoln. Risk assessment of an aging military aircraft. *Journal of Aircraft*, 22(8):687–691, August 1985.

36. J. G. Burns, W. P. Johnson, and A. P. Berens. Aging aircraft structural damage analysis. In *AGARD Conference Proceedings 506*, pages 4.1–4.15, May 1991.

37. J. N. Yang and S. D. Manning. Demonstration of proabilistic based durability analysis method for metallic airframes. *Journal of Aircraft*, 27(2):169–175, 1990.

38. J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods*. Wiley, New York, 1964.

39. NASA. Fracture mechanics reliability guideline No. GD-AP-2304, October 1995.

40. S. D. Manning, J. N. Yang, and J. L. Rudd. Durability of aircraft structures. In J. W. Provan, editor, *Probabilistic fracture mechanics and reliability*. Dordercht, The Netherlands: Martinus Nijhoff, 1987.

41. S. D. Manning and J. N. Yang. USAF Durability Design Handbook: Guidelines for the analysis and design of durable aircraft structures. Technical Report AFFDL-TR-84-3027, Air Flight Laboratory, Wright Patterson AFB, Feb 1984.

42. J. N. Yang, S. D. Manning, and J. L. Rudd. Durability methods development Volume VII. Technical Report AFFDL-TR-79-3118, U.S. Air Force Flight Dynamics Laboratory, September 1982.

43. R. J. H. Wanhill. Damage tolerance engineering property evaluation of aerospace aluminium alloys with emphasis on fatigue crack growth. Technical Report NLR TP 94177U, National Aerospace Laboratory NLR, 1995.

44. D. Y. Wang. An investigation of initial fatigue quality. In *Design of Fatigue and Fracture Resistant Structures*, number STP 761 in ASTM, pages 191–211. American Society for Testing and Materials, Progress in Flaw Growth and Fracture Toughness Testing, 1982.

45. P. J. Noronha. Fastener hole quality. Technical Report AFFDL-TR-78-206, U. S. Air Force Flight Dynamics Laboratory, December 1978.

46. S. A. Barter. Fatigue crack growth in 7050T7451 aluminium alloy thick section plate with a surface condition simulating some regions of the F/A-18 structure. Technical Report DSTO-TR-1458, Defence Science and Technology Organisation Platforms Sciences Laboratory (Australia), July 2003.

47. Q. Sun L. Molent and A.J. Green. The f/a-18 fatigue crack growth data compendium. Technical Report DSTO-TR-1677, DSTO, February 2005.

48. N. E. Frost and D. S. Dugdale. The propagation of fatigue cracks in sheet specimens. *Journal of the Mechanics and Physics of Solids*, 6:92–110, 1958.

49. F. E. Kiddle. Fatigue endurance, crack sensitivity and nucleation characteristics of structural elements in four aluminium-copper alloys. Technical Report CP 1259, Aeronautical Research Council, London, 1974.

50. T. R. Brussat. Advanced life analysis methods - cracking data survey and ndi assessment for attachment lugs. Technical Report AFWAL-TR-84-3080, Air Force Wright Aeronautical Laboratories, September 1984.

51. A. Teofilo-Bent. *Reliability analysis of degrading uncertain structures with applications to fatigue and fracture under random loading.* PhD thesis, University of Newcastle, Australia, 2003.

52. J. N. Yang, S. D. Manning, W. H. Hsi, and J. L. Rudd. Stochastic crack growth models for applications to aircraft structures. In J. W. Provan, editor, *Probabilistic fracture mechanics and reliability.* Dordercht, The Netherlands: Martinus Nijhoff, 1987.

53. K. Ortiz. *On the stochastic modelling of fatigue crack growth.* PhD thesis, Stanford University, Stanford, California, November 1985.

54. E. Ayaloa-Uraga and T. Moan. System reliability issues of off-shore structures considering fatigue failure and updating based on inspection. In *First International ASRANet Colloquium, Glasgow, Scotland*, July 2002.

55. P. J. E. Forsyth and D. A. Ryder. Some results of the examination of aluminium alloy specimen fracture surfaces. *Metallurgica*, 63:117–124, 1961.

56. P. C. Paris and F. Erdogan. A critical analysis of crack propagation laws. *Journal of Basic Eng.*, (85):528–534, 1963.

57. W. Weibull. A statistical theory of strength of materials. *I.V.A.H*, 1939.

58. C. A. Zapffe and C. O. Worden. Fractographic registrations of fatigue. *Trans. Amer. Soc. Metals*, 43:958–969, 1951.

59. G. R. Irwin. Analysis of stresses and strains near the end of a crack traversing a plate. *Journal of Applied Mechanics*, 1957.

60. D. A. Ryder. unknown. Technical Report Met 288, Royal Aircraft Establishment, 1958.

61. J. C. McMillan and R. M. N. Pelloux. unknown. In *Fatigue crack growth under spectrum loads*, number ASTM STP 415, page 505. American Society for Testing and Materials, 1967.

62. P. C. Paris, M. P. Gomez, and W. P. Anderson. A rational analytic theory of fatigue. In *The trend in Engineering*, number 13, pages 9–14. 1961.

63. W. Elber. The significance of fatigue crack closure. In *Damage Tolerance in Aircraft Structures*, number 486 in Special Technical Publication, pages 230–242. American Society for Testing and Materials, 1971.

64. Y. Murakami, editor. *International Symosium on Fatigue Damage Measurement and Evaluation under Complex Loadings 1991 Fukuoka, Japan, the rainflow method in fatigue: The Tasuo Endo Memorial Volume*. Butterworth-Heinmann, 1992.

65. E. K. Walker. The effect of stress ratio during crack propagation and fatigue for 2024-T3 and 7075-T6 aluminium. In *Effects of Environment and Complex Load History on Fatigue Life*, number STP 462, pages 1–14. ASTM, 1970.

66. R. G. Forman, V. E. Kearney, and R. M. Engle. Numerical analysis of crack propagation in cyclicly-loaded structures. *Journal of Basic Engineering Transactions*, 89(3):459–465, 1967.

67. S. A. Barter, L. Molent, N. Goldsmith, and R. Jones. An experimental evaluation of fatigue crack growth. *Engineering Failure Analysis*, 12(1):99–128, February 2005.

68. S. A. Barter, M. McDonald, and L. Molent. Fleet fatigue life interpretation from full-scale and coupon fatigue tests a simplified approach. In *Proc. 2005 USAF Structural Integrity Program (ASIP) Conference, Memphis, Ten, USA, 29th Nov. to 1st Dec. 2005*, 2005.

69. O. E. Wheeler. Spectrum loading and crack growth. *Journal of Basic Engineering*, 94:181–186, March 1972.

70. J. Willenborg, R. M. Engle, and H. A. Wood. A crack growth retardation model using an effective stress concept. Technical Report AFFDL-TR-71-1, U. S. Air Force Flight Dynamics Laboratory, Wright-Patterson AFB, Ohio, January 1971.

71. J. Lankford, Jr. and D. L. Davidson. Fatigue crack tip plasticity associated with overloads and subsequent cycling. *Journal of Engineering Materials and Technology*, pages 17–23, January 1976.

72. D. P. Rooke and D. J. Cartwright. *Compendium of Stress Intensity Factors*. H.M.S.O, 1976.

73. R. Marissen, K. H. Trautmann, and H. Nowack. The influence of compression loads and of dK/da on the crack propagation under variable amplitude loading. *Engineering Fracture Mechanics*, 19:863–879, 1984.

74. A. U. de Koning. A simple crack closure model for prediction of fatigue crack growth rates under variable-amplitude loading. In *Fracture Mechanics: Thirteenth Conference*, pages 63–85. American Society for Testing and Materials, 1981.

75. J. C. Newman, Jr. A crack closure model for predicting fatigue-crack growth under aircraft spectrum loading. In *Methods and Models for Predicting Fatigue Crack Growth Under Random Loading*, number STP 748 in ASTM, pages 53–84. American Society for Testing and Materials, 1981.

76. S. Suresh. Micromechanisms of fatigue crack growth retardation following overloads. *Engineering Fracture Mechanics*, 18:577–593, 1983.

77. M. N. James, M. N. Pacey, L. W. Wei, and E. A. Patterson. Characterisation of plasticity-induced closure – crack flank contact force versus plastic enclave. *Engineering Fracture Mechanics*, 70:2473–2487, 2003.

78. A. K. Vasudevan, K. Sadananda, and J. Louat. A review of crack closure, fatigue crack threshold and related phenomena. *Materials Science and Engineering*, AI88:1–22, 1994.

79. A. U. de Koning and G. Liefting. Analysis of crack opening behaviour by application of a discretised strip yield model. Technical Report MP 87065 U, NLR, March 1986.

80. G. S. Wang and A. F. Blom. A strip model for fatigue crack growth predictions under general load conditions. *Engineering Fracture Mechanics*, 40(3):507–533, 1991.

81. B. N. Leis and T. P. Forte. Fatigue growth of initially physically short cracks in notched aluminium and steel plates. In *Fracture Mechanics: Thirteenth Conference*, pages 100–124. American Society for Testing and Materials, 1981.

82. U.S. Air Force. *Damage Tolerant Design Handbook*, May 1994.

83. J. M. Barsom. Fatigue crack growth under variable amplitude loading. Number 536 in Special Technical Publication. American Society for Testing and Materials, 1976.

84. J. N. Yang. Statisical crack growth in durability and damage tolerant analyses. In *22nd Structures, Structual Dynamics and Materials Conference Atlanta, Georgia*, April 1981.

85. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C - The art of scientific computing*. Cambridge University Press, 1992.

86. NIST, http://www.itl.nist.gov/div898/handbook/. *NIST/SEMATECH e-Handbook of Statistical Methods*, January 2003.

87. R. B. Abernethy. *Weibull analysis handbook*. Pratt and Whitney Aircraft Group, November 1983.

88. Z. Yosibash. Structural risk assessment in the Israeli Air Force. *Journal of Aircraft*, July–August 1992.

89. John W. Lincoln and Ronald A. Melliere. Economic life determination for military aircraft. *Journal of Aircraft*, 36(5), Sep–Oct 1999.

90. J. N. Yang and W. J. Trapp. Joint aircraft loading/structure response statistics of time to service crack initiation. *Journal of Aircraft*, 13(4):270–278, April 1976.

91. J. N. Yang, S. D. Manning, and W. R. Garver. Durability methods development Volume V. Technical Report AFFDL-TR-79-3118, Air Force Flight Dynamics Laboratory, September 1979.

92. J. N. Yang and S. D. Manning. Distribution of equivalent initial flaw size. In *Proceedings, Annual Reliability and Maintainability Symposium, San Francisco, CA*, pages 112–120, January 1980.

93. K. M. Gruenberg, B. A. Craig, and B. M. Hillberry. A probabilistic method for predicting the variability in fatigue behavior of 7075-T6 aluminium. *AIAA*, (AIAA-98-2055), 1998.

94. M. E. Artely. Variations in crack growth behavior. In *11th Fracture Mechanics Conference*, number STP 677, pages 54–67. ASTM, 1979.

95. J. N. Yang and S. D. Manning. Aircraft fleet maintainance based on structural reliability analysis. *Journal of Aircraft*, 31(2), Mar–Apr 1994.

96. W. Weibull. The effect of size and stress history on fatigue crack initiation and propagation. In *Proceedings of the Crack Propagation Symposium*, pages 271–286. Cranfield, 1961.

97. R. Kugel. A relation between theoretical stress concentration factor and fatigue notch factor deduced from the concept of highly stressed volume. In *Proc American Society of Testing and Materials*, volume 61, pages 732–748, 1961.

98. B. Palmberg, A. F. Blom, and S. Eggwertz. Probabilistic damage tolerance analysis of aircraft structures. In J. W. Provan, editor, *Probabilistic fracture mechanics and reliability*. Dordercht, The Netherlands: Martinus Nijhoff, 1987.

99. J. Coyle, J. M. Grandage, and D. G. Ford. The distribution of fracture toughness data for D6ac steel. Technical report, Aeronautical Research Laboratories, August 1976.

100. K. Erland. Quantifying the benefit of redundant fluorescent penetrant inspection. *Review of progress in quanitative nondestructive evaluation*, 8b:2221–2228, 1989.

101. W. H. Lewis, W. H. Sproat, B. D. Dodd, and J. M. Hamilton. Reliability of non-destructive inspections - final report. Technical Report US Air Force SA-ALC/MME 76-6-38-1, Lockheed Georgia Company, December 1978.

102. A. P. Berens and P. W. Hovey. Evaluation of NDE charactisation Volume I. Technical Report AFWAL-TR-81-4160, University of Dayton, Research Institute, Ohio, 1981.

103. J. N. Yang and S. Chen. An exploratory study of retirement for cause for gas turbine engine components. In *Proceedings AIAA/ASME/SAE 20th Joint Propulsion Conference*, number 84-1220, June 1984.

104. A. O. Payne. A reliability approach to the fatigue of structures. In *Probabilistic aspects of fatigue*, number 511 in STP. American Society for Testing and Materials, 1972.

105. A. O. Payne, G. D. Mallinson, and A. D. Graham. Reliability approach to structural safety and safety criteria. In S. Eggwertz and N. C. Lind, editors, *Probabilistic methods in the mechanics of solids and structures IUTAM Symposium Stockholm*, Berlin, 1984. Springer-Verlag.

106. Aircraft Structures. Joint Service Specification Guide JSSG-2006, (US) Department of Defense, October 1998.

107. J. W. Lincoln. Risk assessment of an aging military aircraft. In *ASME, and AHS, Structures, Structural Dynamics and Materials Conference, 24th, Palm Springs, CA, May 14-16, 1984*, number AIAA-1984-851, 1984.

108. G. S. Wang and A. F. Blom. Analytical method to predict initiation of multiple site damage at mechanical joints. In A. P. S. Selvadurai and C. A. Brebbia, editors, *Damage and Fracture Mechanics VI*. WIT Press, 2000.

109. C. Babish, R. Alford, and M. Thomsen. Probabilistic approach to support force management decision making process. In *2003 ASIP Conference*, Savannah Georgia, December 2003.

110. J. W. Lincoln. Aging Aircraft - USAF experiences and actions. In R. Cook and P. Poole, editors, *ICAF97 - Fatigue in new and aging aircraft, Proceedings of the Nineteenth symposium of the international committee on aeronautical fatigue, Edinbourough Scotland, 18–20 June 1997*, 1997.

111. Structural evaluation of the E-8C JSTARS wing lower surface. Technical Interchange Meeting Presentation, Boeing Product Support Division, Wichita, Kansas, April 1996.

112. P. White and B. Dixon. Structural Risk Assessment of the B707C WS360 upper wing splice joint for RAAF aircraft A20-629. Technical report, Defence Science and Technology Organisation, May 2004.

113. P. Jackson, J. Agius, and D. Sexton. Structural airworthiness and cost of ownership implications of continues operation of RAAF B707 aircraft – final report. Technical Report DSTO-TN-0267, Defence Science and Technology Organisation, August 1999.

114. J. W. Lincoln and R. A. Melliere. Economic life determination for a military aircraft. In *AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit, 39th, and AIAA/ASME/AHS Adaptive Structures Forum, Long Beach, CA, Apr. 20-23, 1998*, number AIAA-1998-1828, 1998.

115. J. Branger. Life estimation of a fighter aircraft. In Alfred M. Freudenthal, editor, *International Conference on Structural Safety and Reliability*, pages 213–237. Pergamon Press, April 1969.

116. E. J. Gumbel. *Statistics of Extremes.* Columbia University Press, 1958.

117. J. N. Yang and W. J. Trapp. Reliability analysis of fatigue-sensitive aircraft structures under random loading and periodic inspection. Technical Report AFML-TR-74-29, Air Force Materials Laboratory, February 1974.

118. J. W. Lincoln. Method for computation of structural failure probability for an aircraft. Technical Report ASD-TR-80-5035 AD-A-098-294, United States. Air Force. Systems Command, Aeronautical Systems Division, December 1980.

119. A. P. Berens, J. P. Gallagher, and S. Dhar. Risk analysis for modeling damage at multiple sites. Technical Report TR-94-14, Flight Dynamics Directorate, Wright Laboratory, Air Force Materiel Command, Wright-Patterson AFB, OH, February 1994.

120. T. Y. Torng. B-1B aging aircraft risk assesment using a robust importance sampling method. In *Structures, Structural Dynamics and Materials Conference unknown*, April 1999.

121. M. Burchill and K. C. Watters. Probabilistic life analysis of an F-111 wing based on teardown data. In *Proceedings of an International Conference on Structural Integrity and Fracture, SIF 2004, Brisbane, Australia*, September 2004.

122. P. Hoffman and M. Modarres. P-3 reliability and risk analysis. Presentation to the P-3 SLAP TIM, Silver Legacy Hotel, Reno NV, 23-24 June 2004.

123. E. S. Meyer, S. S. Fields, and P. A. Reid. Projecting aircraft fleet reliability. In *Aging Aircraft Conference*, Orlando, Florida, 2001.

124. L. F. Impellizzeri, A. D. Seigel, and R. A. McGinnis. Evaluation of structural reliabilty analysis procedures as applied to a fighter aircraft. Technical Report AFML-TR-73-150, Air Force Materials Laboratory, September 1973.

125. L. Molent, G. Wallace, and A. Currie. Crack growth and repair repair of multi-site damage of fuselage lap joints. Technical Report ARL-STR-TM-534, DSTO, February 1990.

126. R. Jones, N. Bridgford, L. Molent, and G. Wallace. Bonded repair of multi-site damage. In S. G. Sampath S. N. Atluri and P. Tong, editors, *Structural Integrity of Aging Airplanes*, Berlin/Hieldelburg, 1991. Verlag.

127. T. Swift. Widespread fatigue damage monitoring issues and concerns. In *FAA/NASA International Symposium on Advanced Structural Integrity Methods for Airframe Durability and Damage Tolerance*, June 1994.

128. B. L. Smith, A. L. Hijazi, and R. Y. Miyose. Strength of 7075-T6 and 2024-T3 aluminium panels with multiple-site damage. *Journal of Aircraft*, 29(2), March-April 2002.

DISTRIBUTION LIST

Review of Methods and Approaches for the Structural Risk Assessment of Aircraft

Paul White

	Number of Copies
DEFENCE ORGANISATION	
Task Sponsor	
WGCDR J. Agius	1 (printed)
S&T Program	
Chief Defence Scientist	1
Deputy Chief Defence Scientist Policy	1
AS Science Corporate Management	1
Director General Science Policy Development	1
Counsellor, Defence Science, London	Doc Data Sheet
Counsellor, Defence Science, Washington	Doc Data Sheet
Scientific Adviser to MRDC, Thailand	Doc Data Sheet
Scientific Adviser Joint	1
Navy Scientific Adviser	Doc Data Sheet and Dist List
Scientific Adviser, Army	Doc Data Sheet and Dist List
Air Force Scientific Adviser	Doc Data Sheet
Scientific Adviser to the DMO	Doc Data Sheet and Dist List
Deputy Chief Defence Scientist Platform and Human Systems	Doc Data Sheet and Exec Summ
Chief, AVD	Doc Data Sheet and Dist List
RLAS, AVD	Doc Data Sheet and Dist List
Task Manager: K. Watters	1
Author: P. White	1 (printed)
L. Molent	1 (printed)
S. Barter	1 (printed)
R. Antonio	1
C. Tong	1
DSTO Library and Archives	
Library, Fishermans Bend	1 (printed)
Library, Edinburgh	1 (printed)
Defence Archives	1 (printed)

Capability Development Group

Director General Maritime Development	Doc Data Sheet
Director General Capability and Plans	Doc Data Sheet
Assistant Secretary Investment Analysis	Doc Data Sheet
Director Capability Plans and Programming	Doc Data Sheet
Director General Australian Defence Simulation Office	Doc Data Sheet

Chief Information Officer Group

Head Information Capability Management Division	Doc Data Sheet
AS Information Strategy and Futures	Doc Data Sheet
Director General Information Services	Doc Data Sheet

Strategy Group

Assistant Secretary Strategic Planning	Doc Data Sheet
Assistant Secretary Governance and Counter-Proliferation	Doc Data Sheet

Navy

Deputy Director (Operations) Maritime Operational Analysis Centre, Building 89/90, Garden Island, Sydney Deputy Director (Analysis) Maritime Operational Analysis Centre, Building 89/90, Garden Island, Sydney	Doc Data Sheet and Dist List
Director General Navy Capability, Performance and Plans, Navy Headquarters	Doc Data Sheet
Director General Navy Strategic Policy and Futures, Navy Headquarters	Doc Data Sheet

Air Force

SO (Science), Headquarters Air Combat Group, RAAF Base, Williamtown	Doc Data Sheet and Exec Summ

Army

ABCA National Standardisation Officer, Land Warfare Development Sector, Puckapunyal	Doc Data Sheet (pdf format)
J86 (TCS Group), DJFHQ(L), Enoggera QLD	Doc Data Sheet
SO (Science), Land Headquarters (LHQ), Victoria Barracks, NSW	Doc Data Sheet and Exec Summ
SO (Science), Special Operations Command (SOCOMD), R5-SB-15, Russell Offices Canberra	Doc Data Sheet and Exec Summ
SO (Science), DJFHQ(L), Enoggera QLD	Doc Data Sheet

Joint Operations Command

Director General Joint Operations	Doc Data Sheet
Chief of Staff Headquarters Joint Operation Command	Doc Data Sheet
Commandant, ADF Warfare Centre	Doc Data Sheet

Director General Strategic Logistics	Doc Data Sheet
Intelligence and Security Group	
Assistant Secretary, Concepts, Capabilities and Resources	1
DGSTA, DIO	1
Manager, Information Centre, DIO	1
Director Advanced Capabilities, DIGO	Doc Data Sheet
Defence Materiel Organisation	
Deputy CEO, DMO	Doc Data Sheet
Head Aerospace Systems Division	Doc Data Sheet
Head Maritime Systems Division	Doc Data Sheet
Program Manager Air Warfare Destroyer	Doc Data Sheet
Guided Weapon and Explosive Ordnance Branch GWEO	Doc Data Sheet
CDR Joint Logistics Command	Doc Data Sheet

OTHER ORGANISATIONS

National Library of Australia	1
NASA (Canberra)	1

UNIVERSITIES AND COLLEGES

Australian Defence Force Academy Library	1
Head of Aerospace and Mechanical Engineering, ADFA	1
Hargrave Library, Monash University	Doc Data Sheet

INTERNATIONAL DEFENCE INFORMATION CENTRES

US - Defense Technical Information Center	1
UK - Dstl Knowledge Services	1
Canada - Defence Research Directorate R&D Knowledge and Information Management (DRDKIM)	1
NZ - Defence Information Centre	1

ABSTRACTING AND INFORMATION ORGANISATIONS

Library, Chemical Abstracts Reference Service	1
Engineering Societies Library, USA	1
Materials Information, Cambridge Scientific Abstracts, USA	1
Documents Librarian, The Center for Research Libraries, USA	1

INFORMATION EXCHANGE AGREEMENT PARTNERS

National Aerospace Laboratory, Japan	1

National Aerospace Laboratory, Netherlands	1

SPARES

DSTO Edinburgh Library	5 (printed)

Total number of copies: printed 12, pdf 25

Page classification:UNCLASSIFIED

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. CAVEAT/PRIVACY MARKING

2. TITLE	3. SECURITY CLASSIFICATION
Review of Methods and Approaches for the Structural Risk Assessment of Aircraft	Document (U) Title (U) Abstract (U)
4. AUTHORS	5. CORPORATE AUTHOR
Paul White	Defence Science and Technology Organisation 506 Lorimer St, Fishermans Bend, Victoria 3207, Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO–TR–1916	013–746	Technical Report	October, 2006

8. FILE NUMBER	9. TASK NUMBER	10. SPONSOR	11. No OF PAGES	12. No OF REFS
2005/1078938	AIR 03/192	ASI	65	128

13. URL OF ELECTRONIC VERSION	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO–TR–1916.pdf	Chief, Air Vehicles Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved For Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DSTO RESEARCH LIBRARY THESAURUS

Aircraft structures, Probability theory, Service life, Reliability engineering

19. ABSTRACT

This report looks at the published literature on methods and assumptions made in performing structural risk assessments on aircraft. Because the major contributor to the risk of structural failure is fatigue, most methods of risk assessment involve modelling the effect of fatigue growth by some probabilistic method. Many risk assessments use the equivalent initial flaw size approach to allow for the variability in fatigue crack growth. Common errors in the formulation are made in many risk assessments, which can be significant and are described in this report. It is found that the standard approach can produce an acceptable assessment of the probability of failure of an aircraft if care is taken in understanding what is being modelled and the assumptions on which the analysis is based. A number of case studies of risk assessment's performed on different aircraft are summarised.

Page classification:UNCLASSIFIED